

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Nov 5 09:13:43 1998; MasPar time 2713.32 Seconds
Tabular output not generated. 1526.778 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGCGGCGCGC 2027
Comp: CTTAAGCCGGAAGTACCGGA.....TTTTTTTTTTTCGCGCGCGC

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 552174 seqs, 1021863385 bases x 2

Post-processing: Minimum Match 0%
Listing first 60 summaries

Database: emb155
1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro
genbank107
Database: 12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
18:gb_ph 19:gb_pl 20:gb_pr1 21:gb_pr2 22:gb_ro 23:gb_st
24:gb_sts 25:gb_sy 26:gb_un 27:gb_v1

Statistics: Mean 12.049; Variance 7.022; scale 1.716

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1043	51.5	2028	22	MMU88567 Mus musculus secreted	0.00e+00
2	993	49.0	2031	22	AF017989 Mus musculus secreted	0.00e+00
3	959	47.3	1776	22	D50462 Mouse SDF5 mRNA, compl	0.00e+00
4	766	37.8	882	21	AF017986 Homo sapiens secreted	0.00e+00
5	110	5.4	1984	21	AF017988 Homo sapiens secreted	8.64e-50
6	106	5.2	2094	21	AF017987 Homo sapiens secreted	4.01e-47
7	106	5.2	4240	15	BTU85945 Bos taurus frizzled-re	4.01e-47
8	103	5.1	2075	21	AF001900 Homo sapiens secreted	3.93e-45
9	103	5.1	4469	21	AF056087 Homo sapiens secreted	3.93e-45
10	101	5.0	2659	22	MMU88566 Mus musculus secreted	8.29e-44
11	97	4.8	7218	17	I66494 Sequence 14 from paten	3.60e-41
12	96	4.7	7218	17	I66494 Sequence 14 from paten	1.63e-40
13	83	4.1	2075	16	AF006508 Gallus gallus crescent	4.64e-37
14	60	3.0	2184	21	HSU82169 Human frizzled homolog	1.15e-17
15	60	3.0	106049	13	AC005049 Homo sapiens clone RG0	1.15e-17

16	3.0	136389	13	AC005074	Homo sapiens clone RG2	1.15e-17
17	2.9	2334	20	HSU43318	Human putative transme	1.84e-16
18	2.8	2328	16	AF031830	Gallus gallus 7-transm	7.29e-16
19	2.8	2421	22	MMU43321	Mus musculus putative	2.87e-15
20	2.5	1881	22	AF054623	Mus musculus frizzled-	2.49e-12
21	2.4	1181	16	AF059570	Xenopus laevis secrete	1.32e-10
22	2.4	1851	16	AF017177	Xenopus laevis frizzle	3.55e-11
23	2.4	10772	14	AF012089	Drosophila melanogaste	3.55e-11
24	2.3	10772	14	AF012089	Drosophila melanogaste	1.79e-09
25	2.3	74371	21	AC005369	Homo sapiens chromosom	1.79e-09
26	2.2	1912	22	RATFRZ2H	Rattus norvegicus Dros	6.52e-09
27	2.1	1727	16	AF031831	Gallus gallus 7-transm	2.95e-07
28	2.0	1923	21	HUMFRIZ	Human frizzled gene pr	1.03e-06
29	1.9	215	17	I28278	Sequence 5 from patent	1.22e-05
30	1.9	4540	22	RATFRZ2H	Rattus norvegicus Dros	4.10e-05
31	1.7	74371	21	AC005369	Homo sapiens chromosom	4.62e-03
32	1.7	216021	21	HUAC004787	Homo sapiens Chromosom	4.62e-03
33	1.5	215	17	I28278	Sequence 5 from patent	1.35e-01
34	1.5	354	14	OFU89259	Oxytricha fallax 57kd	4.02e-01
35	1.5	4297	19	ATHSC701	A.thaliana hsc70-1 gen	4.02e-01
36	1.5	38262	14	CELK02A2	Caenorhabditis elegans	1.35e-01
37	1.5	147898	13	HS1111N9	Human DNA sequence ***	1.35e-01
38	1.4	156	24	DM36D7S	D. melanogaster STS de	1.17e+00
39	1.4	307	14	AF010529	Plasmodium falciparum	1.17e+00
40	1.4	400	24	G13598	human STS SHGC-11851.	3.33e+00
41	1.4	1073	19	ATCOR47	A.thaliana cor47 mRNA.	1.17e+00
42	1.4	1461	22	AF045766	Mus musculus orphan G	3.33e+00
43	1.4	1466	17	A07367	M.musculus VAC-alpha c	3.33e+00
44	1.4	1566	21	HUMBCI	Homo sapiens mRNA for	3.33e+00
45	1.4	1566	17	E03096	DNA encoding calphobin	3.33e+00
46	1.4	1567	17	I33410	Sequence 2 from patent	3.33e+00
47	1.4	1575	17	AI3845	H.sapiens PP4 mRNA.	3.33e+00
48	1.4	1582	21	HUMPAP4	Human placenta antioxa	3.33e+00
49	1.4	1592	20	HUMENN	Human endonexin II mRN	3.33e+00
50	1.4	1605	17	I07181	Sequence 12 from Paten	3.33e+00
51	1.4	2028	21	HSU56813	Human polycystwin mRNA	1.17e+00
52	1.4	2112	22	MMEBK	M.musculus mRNA for Eb	1.17e+00
53	1.4	2260	22	MMU43320	Mus musculus putative	1.17e+00
54	1.4	4798	14	PFSC04076	Plasmodium falciparum	1.17e+00
55	1.4	7964	14	PFSC04015	Plasmodium falciparum	1.17e+00
56	1.4	8595	14	DDU14576	Dictyostelium discoide	1.17e+00
57	1.4	10934	12	ECAE000479	Escherichia coli K-12	3.33e+00
58	1.4	110102	19	ATF28J12	Arabidopsis thaliana D	3.33e+00
59	1.4	115641	19	F8K4	Arabidopsis thaliana c	1.17e+00
60	1.4	176195	12	ECOUW89	E. coli Chromosomal re	3.33e+00

ALIGNMENTS

RESULT	1	MMU88567	2028 bp	mRNA	ROD	22-APR-1997
LOCUS		Mus musculus secreted	frizzled related protein	SFRP-2	(Sfrp2) mRNA,	
DEFINITION		complete cds.				
ACCESSION		U88567				
NID		g1946342				
KEYWORDS						
SOURCE		house mouse.				
ORGANISM		Mus musculus				
REFERENCE						
AUTHORS		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE		1 (bases 1 to 2028)				
JOURNAL		Rattner, A., Hsieh, J.-C., Smallwood, P.M., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.				
MEDLINE		A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors				
REFERENCE		Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)				
AUTHORS		2 (bases 1 to 2028)				
TITLE		Rattner, A., Hsieh, J.-C., Smallwood, P.M., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.				
JOURNAL		Direct Submission				
MEDLINE		Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns				

Hopkins University School of Medicine, 725 North Wolfe Street, PCTB
805, Baltimore, MD 21205

FEATURES
source
1. .2028
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="3"
/map="near F9g"
236. .1123
/gene="Sfrp2"
236. .1123
/gene="Sfrp2"
/codon_start=1
/product="secreted frizzled related protein SFRP-2"
/db_xref="PID:gl946343"
/translation="MPRGPASL L L L V L A S H C C L G S A R G L F L F G Q P D F S Y K R S N C K P I P
A N L Q L C H G I E Y Q N M R L P N L L G H E T M K E V L E Q A G A W I P L V M K Q C H P D T K K F L C S L F A P V
C L D D L D E T I Q P C H S L C V Q V K D R C A P V M S A F G P W P M L E C D R F P Q D N D L C I P L A S S D H
L L P A T E E A P K V C E A C K T N E D D N D I M E T L C K N D F A L K I K V E I T Y I N R D T K I I L E T K S
K T I Y K L N G V S E R D L K S V L W L K D S L Q C T C E E M N D I N A P Y L V M G Q K Q G G E L V I T S V K R W
Q K G Q R E F K R I S R I R K L Q C"

BASE COUNT 469 a 586 c 480 g 492 t 1 others
ORIGIN

Query Match 51.5%; Score 1043; DB 22; Length 2028;
Best Local Similarity 81.8%; Pred. No. 0.00e+00;
Matches 1558; Conservative 0; Mismatches 312; Indels 34; Gaps 24;

Db 127 CCAGAGCCCCACGAGCAGAGCGGAGGAGTCCCGGACGAGCTCGAGCTCCGGCCGCTCT 186
QY 147 CCGAGTGCACCCAGCGAAGAGAGCGGGCCCCGGACAAGCTCGAACTCCGGCCGCTCG 206
Db 187 CGCTTCCCGCTCGGCTCCCTCCGCCCGCCCGGGGTGCGTAGTCCACGATGCCGCGGG 246
QY 207 CCCTTCCCGGCTCCGCTCCCTCTGCCCTCGGGTCCGCCGCGCACGATGCTGCAGG 266
Db 247 CCCTGCCCTCGCTGCTGCTAGTCTCGCTCGCTCGCACTGCTGCCGTGGCTCGGCGGTG 306
QY 267 CCCTGGCTCGCTGCTGCTCTCTCTCGCTCGCACTGCTGCCGTGGCTCGGCGCGCG 326
Db 307 GCTCTTCCCTCTTCGGCCAGCCCGACTTCTCTACAAAGCGCAGCAACTGCAAGCCCATCCC 366
QY 327 GCTCTTCCCTCTTTGGCCAGCCCGACTTCTCTACAAAGCGCAGCAATTGCAAGCCCATCCC 386
Db 367 CGCCAACTGCAGCTGTGCCACGGCATCGAGTACCAGAACATCGGGCTGCCCAACTGCT 426
QY 387 GGCCAACTGCAGCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACTGCT 446
Db 427 GGGCCACGAGACCATGAAGGAGTGCTGGAGCAGGCGGCGCTGGATTCGGTGGTCA 486
QY 447 GGGCCACGAGACCATGAAGGAGTGCTGGAGCAGGCGGCGCTGGATTCGGTGGTCA 506
Db 487 GAAACAGTGCCACCCGGACACCAAGAAGTTCCCTGTGCTCGCTCTTCGCCCTGTCTGTCT 546
QY 507 GAAGCAGTGCCACCCGGACACCAAGAAGTTCCCTGTGCTCGCTCTTCGCCCTGTGCTG 566
Db 547 CGACGACCTAGATGAGACCATCCAGCCCGTGTCACTCGCTCTCGGTGCAGGTGAAGGACCG 606
QY 567 CGATGACCTAGACGAGACCATCCAGCCATGCCACTCGCTCGCTCTCGGTGCAGGTGAAGGACCG 626
Db 607 CTGCGCCCGGTCATGTCCGCTTCGGCTTCGCCCTGGCCAGACATGCTGGAGTCCGACCG 666
QY 627 CTGCGCCCGGTCATGTCCGCTTCGGCTTCGCCCTGGCCCGACATGCTTGAGTCCGACCG 686
Db 667 TTTCCGCGAGGACCAACGACCTCTGCATCCCGCTCGCTAGTAGGACCACTCCTGCCGCG 726
QY 687 TTTCCCGCGAGGACCAACGACCTTGACATCCCGCTCGCTAGCAGCGACCACTCCTGCCAGC 746
Db 727 CACAGAGGAAGCTCCCAAGGTGTGTGAAGCCTGCAAAACCAAGAATGAGGACGACAACGA 786
QY 747 CACCGAGGAAGCTCCCAAGGTATGTGAAGCCTGCAAAACCAAGAATGATGATGACAACGA 806
Db 787 CATCATGGAAACCCCTTTGTAAAAATGACTTCGCACTGAAATCAAAAGTGAAGGAGATAAC 846

QY 807 CATAATGGAACCGCTTTGTAAAAATGATTTGCACTGAAAAATAAAAGTGAAGGAGATAAC 866
Db 847 GTACATCAACAGAGACACCAAGATCATCTCTGGAGACAAGAGCAAGACCATTTACAAGCT 906
QY 867 CTACATCAACCGAGATACCAAAATCATCTGGAGACCAAGAGCAAGACCATTTACAAGCT 926
Db 907 GAACGGCTGTCCGAAAAGGACCTGAAGAAATCCGTGCTGTGGTCAAGACAGCCTGCA 966
QY 927 GAACGGTGTGTCCGAAAAGGACCTGAAGAAATCGGTGCTGTGGCTCAAGACAGCTTGA 986
Db 967 GTGCACCTGTGAGGAGATGAACGACATCAACGCTCCGTATCTGGTCAATGGGACAGAAGCA 1026
QY 987 GTGCACCTGTGAGGAGATGAACGACATCAACGCGCCCTATCTGGTCAATGGGACAGAAACA 1046
Db 1027 GGGCGGCGAGCTGGTGATCACCTCCGTGAAACGGTGGCAGAAGGGCCAGAGAGAGTTCAA 1086
QY 1047 GGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGAAGGGGACAGAGAGTTCAA 1106
Db 1087 GCGCATCTCCCGCAGCATCCGCAAGCTCGAATGCTAGTTTCC-CAGTGGGTGGCTTCTCT 1145
QY 1107 GCGCATCTCCCGCAGCATCCGCAAGCTCGAGTCTAGTCCCGGCATCCTGATGGCTCCGA 1166
Db 1146 CCATCCAGGCGCTGAGCTCTGTAGACCACTCCGCTCCGGCACCTCATTTCCGGTTTCCC 1205
QY 1167 CAGGCGCTGCTCCAGAGCAGCGGTGACCATTTCTGCTCCGGGATCTCAGCTCCCGTTCCC 1226
Db 1206 AAGCACAGTCCGGGAAAGCTACAGCCCGCAGCTTGGAGCCGCTTGCCCTGCCCTCCTGCATG 1265
QY 1227 AAGCACACTCTTAGCT-GCTCCAGTCTCAGCCTGGCAGCTTCCCCCTGCCCTTTTGCAGG 1285
Db 1266 TGTGTATCCCTAACATGTCTGAGTTATTAAGGCCCTAGGAGGCTTGGAAAACCCATAGCTG 1325
QY 1286 TTTGCATCCCGCAGCATTTCTGAGTTATTAAGGCCACAGGAG---TGA-----TAGCTG 1336
Db 1326 TTTTCACGGAAGCGAAAAGCCCATCCAGATCTTGACAAATATTCAAACATAATAAAATC 1385
QY 1337 TTTTCACCTAAAG-GAAAAGCCCAACCGAATCTTGAGAAATATTCAAACATAATAAAATC 1395
Db 1386 ATGACTATTTTATGAAGTTTATAGACAGCTCGTTTAAAGTTAGTTTGAATAGCTGTA 1445
QY 1396 ATGAATATTTTATGAAGTTTAAATAGCTCAGTTTAAAGCTAGTTTGAATAGGTGCA 1455
Db 1446 GTACTTTGACCCGAGGGCATTTTCTCTCTTTGGTCAGTCTGTGGCTTATACCGTGCACT 1505
QY 1456 --ACTGTGACTTG-GGTCTGGTTGGTTGTT--GTTTGTGTTTGTAGTC-AGC-TGATTT 1508
Db 1506 TAGGTTGCCATGTGAGGCGAATTTGTTCTTTTTTTTTTTTTTTCTCTGTGGTCTAA 1565
QY 1509 TCAGTTCCAC-TGAGGTTGTCAACATGCAAAATGCTTCAATTTTCTGTGGCCCAA 1567
Db 1566 GCTTGTGGNCCCAGACTTAGTTGAGATAAAGCTGGCTGTTATCTCAAAGTCTTCCCTCAG 1625
QY 1568 ACTTGTGGTCAAAAACCCCTGTTGAGATAAAGCTGGCTGTTATCTCAACATCTTTCATCAG 1627
Db 1626 TTCCAGGCGCTGAGAATCGGCATCTAAGTCTTCAAACATTTTCGTTGCTCGTTTATGCCCT 1685
QY 1628 CTCCAGAC-TGAGACTCAGTGTCTAAGTCTTACAACAATTCATCATTTTATACCTTCAAT 1686
Db 1686 CATGAGCTCTGACCATTTGCATGCGTTCCCATCCAGCTACAGAACTTCAGTTTATAAGCA 1745
QY 1687 GG-GAAGCTAAACTGTTACATGTAT-CACATTCAGCTACATACTTCCATTTATTAGAA 1744
Db 1746 -CACAGTAACCATTCCTCATTTGATGATGCCCTCAAATAAAAAGTGAATACAGTCTATAA 1804
QY 1745 GCACATTAACCATTTCT-ATAGCATGATTTCTTCAAGTAAAAGGCAAA-AGA-TATA-AA 1800
Db 1805 ATTGACGAGTATTTTAAGCTTTGTTAAAAACATCTTTTAAATCAATTTTAAATCAATTTT 1864
QY 1801 TTTTATAAATGACTTGAG--TACTTTAAGCCTTGTTTAAAA-CA-TTCTTACTTAACCTT 1856
Db 1865 TTTTGCAAACTAAATCATTTGAGCTTACCTGTAATATACGTAAGTAGTTGACCTGGAAAAG 1924

QY 1857 TTGCA-AATTAAACCCATTGTAGCTTACCTGTAATATACATAGTAGTTTACCTTTAAAG 1915

Db 1925 TTGTAAAAATATGCTTTAAACGACACTGTAATATTTTACAGATAAACATTATATCTTTG 1984

QY 1916 TTGTAAAAATATGCTTTAAACCAACACTGTAATATTTTACAGATAAACATTATATCTT-G 1974

Db 1985 TATATAAACTTTACATCTTGTGTACCTATATAAAAAAAAAAAAA 2028

QY 1975 TATATAAACTTTACATCTTGTGTACCTATATAAAAAAAAAAAAA 2018

RESULT 2

LOCUS AF017989 2031 bp mRNA ROD 21-SEP-1997

DEFINITION Mus musculus secreted apoptosis related protein 1' (Sarpl) mRNA, complete cds.

ACCESSION AF017989

NID 92415420

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2031)

AUTHORS Melkonyan,H., Chang,W.C., Shapiro,J.P., Mahadevappa,M., Fitzpatrick,P.A., Kiefer,M.C., Tomei,D.L. and Umansky,S.R.

TITLE SARPs - a new family of proteins that regulate apoptosis

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2031)

AUTHORS Melkonyan,H., Prochazka,V. and Umansky,S.R.

TITLE Direct Submission

JOURNAL Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way South, Richmond, CA 94804, USA

FEATURES

source Location/Qualifiers

1..2031

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="10T1/2 quiescent cells"

/dev_stage="embryo"

1..2031

/gene="Sarpl"

254..1141

/gene="Sarpl"

/note="similar to frizzled-like receptor CRD"

/codon_start=1

/product="secreted apoptosis related protein 1"

/db_xref="PID:g2415421"

/translation="MPRGPASLLLLVLASHCLGSARGFLFGOPDFSYKRTNCKPIP ANLQLCHGIEYQNMRLPNLGHETMKEVLEQAGAWIPLVMKQHPDTKKFLCSLFAPV CLDDLDETIQPCHSVCVQKDRCAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDH LLPATEEAPKVCEAKTKNEDNDIMETLCKNDFALKIKVKEITYINRDKIILETKS KTIYKLVNGVSRDLKSVLWLDLSLQCTCEEMNDINAPLYVMQKQGQGGELVITSVKRW QKGQREFKRISRSIRKLQC"

BASE COUNT 460 a 591 c 485 g 495 t

ORIGIN

Query Match 49.0%; Score 993; DB 22; Length 2031;

Best Local Similarity 81.3%; Pred. No. 0.00e+00;

Matches 1522; Conservative 0; Mismatches 313; Indels 36; Gaps 25;

Db 145 CCAGAGCCCCACGACGAGAGCGAGGGAGTCCCGACGAGCTCGAGCTCCGGCGCCTCT 204

QY 147 CCCAGTCCGACCCAGCGAAGAGAGCGGGCCCGGACAAAGCTCGAACTCCGGCGCCTCG 206

Db 205 CGCTTCCCGCTCGGCTCCCTCCGCCCCCGGGGGTTCGCTAGTCCACGATCGCGCGGG 264

QY 207 CCCTTCCCGCTCCGCTCCCTCTGCCCTCGGGTTCGGCGCCACGATGCTGCAGGG 266

Db 265 CCCTGCTCGCTGCTGCTAGTCTCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 324

QY 267 CCCTGGCTCGCTGCTGCTCTTCTTCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 326

Db 325 GCTCTCTCTTTCGGCCAGCCCGACTTCTCTCTACAGCGCACGAAGCTGCAAGCCCATCCC 384

QY 327 GCTCTTCTCTTTGGCCAGCCCGACTTCTCTACAAAGCGCAGCAATTGCAAGCCCATCCC 386

Db 385 CGCCAAACCTGCAGCTGTGCACGGCATCGAGTACCAGAACATCGGGTGCCTCCAAACCTGCT 444

QY 387 GGCCAAACCTGCAGCTGTGCACGGCATCGAATACCAGAACATCGGGTGCCTCCAAACCTGCT 446

Db 445 GGGCCACGAGACCATGAAGAGGTGCTGGAGCAGGCGGCGCTTGGATTCCGCTGGTCTAT 504

QY 447 GGGCCACGAGACCATGAAGAGGTGCTGGAGCAGGCGGCGCTTGGATTCCGCTGGTCTAT 506

Db 505 GAAGCAGTGCCACCCGGACACCAAGAAGTTCTCTGTCTCGCTCTTCGCCCTGTCTGTCT 564

QY 507 GAAGCAGTGCCACCCGGACACCAAGAAGTTCTCTGTCTCGCTCTTCGCCCTGTCTGCT 566

Db 565 CGACGACCTAGATGAGACCATCCAGCCGTGTCACTCGCTCTGCTGCTGCTGCTGCTGCTG 624

QY 567 CGATGACCTAGACGAGACCATCCAGCCATGCCACTCGCTCGCTGCTGCTGCTGCTGCTG 626

Db 625 CTGCGCCCCGGTTCATGTCGGCTTCGGCTTCGCCCTGGCCAGACATGCTGGAGTCCGACCG 684

QY 627 CTGCGCCCCGGTTCATGTCGGCTTCGGCTTCGCCCTGGCCAGACATGCTTGAGTCCGACCG 686

Db 685 TTTCGGCAGGACAAACGACCTCTGCATCCCCCTCGCTAGTAGGACCACTCTCTGCCCGC 744

QY 687 TTTCGGCAGGACAAACGACCTTTGCATCCCCCTCGCTAGCAGCAGCACCTCTCTGCCAGC 746

Db 745 CACAGAGGAAGCTCCCAAGGTGTGTGAAGCTTGCACCTGCAAAACCAAGAATGAGGACGACA 804

QY 747 CACGAGGAAGCTCCCAAGGTGTGTGAAGCTTGCACCTGCAAAACCAAGAATGATGATGACA 806

Db 805 CATCATGGAACCCCTTTGTAAAAAATGACTTCGCACCTGAAATCAAAGTGAAGGAGATAAC 864

QY 807 CATAATGGAACGCTTTGTAAAAAATGATTTTGCACTGAAATAAAAGTGAAGGAGATAAC 866

Db 865 GTACATCAACAGAGACACCAAGATCATCTCTGGAGACAAAGAGCCAGACCATTTTACAAGCT 924

QY 867 CTACATCAACCGAGATACCAAAATCATCTCTGGAGACCAAGAGCAAGACCATTTTACAAGCT 926

Db 925 GAACGGCGTGTCCGAAAGGACCTGAAGAAATCCGTGCTGTGGCTCAAAGACAGCCTGCA 984

QY 927 GAACGGTGTGTCCGAAAGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAGCCTTGCA 986

Db 985 GTGCACCTGTGAGGAGATGAACGACATCAACGCTCCGTAATCTGGTTCATGGGACAGAGCA 1044

QY 987 GTGCACCTGTGAGGAGATGAACGACATCAACGCGCCCTATCTGGTTCATGGGACAGAAACA 1046

Db 1045 GGGCGGGAACCTGGTGATCACCTCCGTGAAACGGTGGCAGAGGGCCAGAGAGAGTTCAA 1104

QY 1047 GGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGAGGGCCAGAGAGAGTTCAA 1106

Db 1105 GCGCATCTCCGCGAGCATCCGAAAGCTGCAATGCTAGTTTCCAGTGGGGTGGCTTCTCT 1164

QY 1107 GCGCATCTCCGCGAGCATCCGAAAGCTGCAGTGTAGTTCGCGGCATCCTGATGGCTCCGA 1166

Db 1165 CCATCCAGGCCCTGAGCTCTGTAGACCACTTGC-CTCCGG-ACCTCATTTCCGGTTTCCC 1222

QY 1167 CAGGCTGCTCCAGAGCAGCGCTGACCATTTCTGCTCCGGGATCTCAGCTCCCGTTCCCC 1226

Db 1223 AAGCACAGTCCGGGAAAGCTACAGCCCCCAGCTTGGAGCCGCTTGCCTCTGCTGCTG 1282

QY 1227 AAGCACACTCCTAGCT-GCTCCAGTCTCAGCCTGGGCGAGCTTCCCCCTTTTGCACG 1285

Db 1283 TGTGTATCCCTAACATGTCCTGAGTTATAAGGCCCTTAGGAGCCCTTGGAAACCCATAGCT 1342

QY 1286 TTGTCATCCCCAGCATTTCTGTAGTTATAAGGCCACAGGAG---TGGA-----TAGCT 1335

Db 1343 GTTTTACGGAAAGCGAAAGCCCATCCAGATCTTGTACAAATATTCAAACATAATAAAAT 1402

QY 1336 GTTTTACCTAAAG-GAAAAGCCACCCCGAATCTTGTAGAAATATTCAAACATAATAAAAT 1394

Db 1403 CATGACTATTTTATGAAGTTTATAGAACAGCTCGTTTTTAAGTTAGTTTGAATAGCTGT 1462

QY 1395 CATGAATATTTTATGAAGTTTAAATAATAGCTCACCTTTAAAGCTAGTTTGAATAGGTGC 1454


```

|||||
QY 911 AGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAAATCGGTGCTGTGGC 970
Db 729 TCAAAGACAGCCTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGCTCCGTAATCTGG 788
QY 971 TCAAAGACAGCTTGCAAGTGCACCTGTGAGGAGATGAACGACATCAACGCGCCCTATCTGG 1030
Db 789 TCATGGGACAGAGCAGCGCGGAGCTGGTGATCACCTCCGTAACCGGTGGCAGAAGG 848
QY 1031 TCATGGGACAGAAAGGGTGGGAGCTGGTGATCACCTCGGTGAACGGTGGCAGAAGG 1090
Db 849 GCCAGAGAGAGTTCAACCGCATCTCCCGCAGCATCCGCAAGCTGCAATGCTAGTTCCCA 908
QY 1091 GGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGTAGTCCCGGC 1150
Db 909 GTGGGGTGGCTTCTCTCCATCCAGGCCCTGAGCTGTAGACCACTTCCGCTCCGGACC 968
QY 1151 ATCCTGATGGCTCCGACAGGCCTGTCTCCAGAGCAGCGCTGACCATTTCTGCTCCGGGATC 1210
Db 969 TCATTTCCGGTTTCCCAAGCACAGTCCGGGAAAGCTACAGCCCCAGCTTGGAGCCGCTTG 1028
QY 1211 TCAGCTCCCGTTTCCCAAGCACACTCCTAGCT-GCTCCAGTCTCAGCCTGGGCAGCTTCC 1269
Db 1029 CCCTGCCCTCCTGCATGTGTATCCCTAACATGCTCCTGAGTTATAAGGCCCTAGGAGGCC 1088
QY 1270 CCCTGCCCTTTGGCAGCTTTGCATCCCCAGCATTTCTGTAGTTATAAGGCCACAGAG- 1326
Db 1089 TTGGAAACCCATAGCTGTTTTTCAGGAAAGCGAAAGCCCATCCAGATCTTGTACAAATA 1148
QY 1327 -TGA-----TAGCTGTTTTCACCTAAAG-GAAAGAGCCACCCGAATCTTGTAGAAATA 1378
Db 1149 TTCAAACTAATAAATCATGACTATTTTATGAGTTTITAGAACAGCTCGTTTAAAGTT 1208
QY 1379 TTCAAACTAATAAATCATGAATATTTTATGAGTTTAAATAAGCTCAGCTTAAAGCT 1438
Db 1209 AGTTTGAATAGCTGTAGTACTTTGACCCGAGSGGSCATTTTCTCTCTTTGGTCAGTCTGT 1268
QY 1439 AGTTTGAATAGTGCA--ACTGTGACTTGGGTCTGGTGGTGTGTTGTTGTTGT-T-T 1493
Db 1269 TGGCTTATACCGTGCACTTAGGTTGCCATGTACAGCGGAATTGTTCTTTTCTTTTCTTTT 1328
QY 1494 TGAGTCAG--C-TGATTTTCACTTCCAC-TGAGGTTGTATACATGCAAAATGCTTCA 1549
Db 1329 TTTCCCTCTGTGCTTAAGCTTGTGGTCCCAGACTTGTAGTATGAGATAAAGCTGGCTGTTA 1388
QY 1550 ATTTTCTGTGGCCCAAACTTGTGGTGCACAAACCCTGTTGAGATAAAGCTGGCTGTTA 1609
Db 1389 TCTCAAAGTCTCTCAGTTCAGCCTGAGAAATCGGCATCTAAGTCTTCAAACATTTCTGT 1448
QY 1610 TCTCAACATCTTCATCAGCTCCAGACTGAGACACAGTGTCTAAGTCTTACAAACATTCAT 1669
Db 1449 TGCTCGTTTTATGCCCTCATGAGCTCTGACCATTGCATGCGTTCCTCCATCCAGCTACAGA 1508
QY 1670 CATTTTATACCTTCAATGG-GAAGTTAACTGTGTACATGTAT-CACATTCAGCTACAAT 1727
Db 1509 ACTTCAGTTTATAAGCA-CACAGTAACCATTCCTCATTTGCATGATGCCCTCAAATAAATA 1567
QY 1728 ACTTCCATTTATAGAAGCACATTAACCATTTCT-ATAGCATGATTTCTTCAAGTAAAG 1786
Db 1568 GTGAATACAGTCTATAAATGACGAGTATTTTAAAGCTTTGTTTAAACATCTTTTAAATTC 1627
QY 1787 GCAA-AGA-TATA-AATTTTATAATTGACTTGAG--TACTTTAAGCCTTGTTTAAAA-C 1840
Db 1628 AATTTTATATCATTTTTTTGCAAACATAAATCATTTAGCTTACCTGTAATATACGTAG 1687
QY 1841 A-TTCTCTACTTAACCTTTTGCA-AATTAACCCATGTAGCTTACCTGTAATATACATAG 1898
Db 1688 TAGTTGACCTGGAAGATTGTAAAAATATTGCTTTAACCAGACACTGTAAATATTTTCAGAT 1747
QY 1899 TAGTTTACCTTTAAAGATTGTAAAAATATTGCTTTAACCACACTGTAAATATTTTCAGAT 1958
Db 1748 AAACATTATATCTTTGTATATAAACTTTT 1776
|||||
```

```

QY 1959 AACATTATATTCTT-GTATATAAACTTT 1986

RESULT 4 AF017986 882 bp mRNA PRI 21-SEP-1997
LOCUS Homo sapiens secreted apoptosis related protein 1 (SARP1) mRNA,
DEFINITION partial cds.
ACCESSION AF017986
NID g2415414
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 882)
Meikonyan,H., Chang,W.C., Shapiro,J.P., Mahadevappa,M.,
Fitzpatrick,P.A., Kiefer,M.C., Tomei,D.L. and Umansky,S.R.
SARPs - a new family of proteins that regulate apoptosis
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 882)
AUTHORS Meikonyan,H., Prochazka,V. and Umansky,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
South, Richmond, CA 94804, USA
FEATURES
source
1. .882
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
1. .>882
/gene="SARP1"
265. .>882
/gene="SARP1"
/note="similar to frizzled-like receptor CRD"
/codon_start=1
/product="secreted apoptosis related protein 1"
/db_xref="PID:g2415415"
/translation="MLQPGPSLLLLFLASHCCLSGARGLFLGQPDFSYKRSNCKPIP
ANLQLCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLEAPV
CLDDLDETIQCHSRVQVKDRCAPVMSAFPWPDMLECDRFPQDNDLCIPLASSDHL
PATEEAPKVCCEAKNKNDNDIMETLCKNDFALKIKVKEITYINR"
BASE COUNT 167 a 318 c 233 g 163 t 1 others
ORIGIN
```

Query Match 37.8%; Score 766; DB 21; Length 882;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 841; Conservative 0; Mismatches 10; Indels 11; Gaps 4;

```

Db 29 CTGGCTCATTTCTGCCCGGGTCGGAGCCCCCGGAGCTGCGCGGGCTTGACGCGC 88
QY 19 CTAGCTCATTTCTGCCCGGGTCGGAGCCCCCGGAGCTGCGCGGGCTTGACGCGC 78
Db 89 CTGCGCGCGCTGTCTCCCGGTGTCCCGCTTCTCCGCGCCCCAGCCGCGGCTGCCAGC 148
QY 79 CTGCGCGCGCT--CCTCCCGGTGTCCCGCTTCTCCGCGCCCCAGCCGCGGCTGCCAGC 136
Db 149 TTTTCGGGGCCCCCGAGTCGACCCAGCGAAGAGAGCGGGCCCCGGGACAAAGCTCGAATCC 208
QY 137 TTTTCGGGGCCCCCGAGTCGACCCAGCGAAGAGAGCGGGCCCCGGGACAAAGCTCGAATCC 196
Db 209 GGCGGCTCGCCCTTAACCAAGCTCCG-TCCCTCTACCCCTAGGGGTGCGCG--CCACGA 265
QY 197 GGCGGCTCGCCCTTCCCGGCTCCGCTCCCTCTGCCCCCTCGGGTGGCGGCCACGA 256
Db 266 TGCTGCAGGGCCCTGGCTCGCTGCTGCTCTCTCTCCCTCGCCTCGCACTGCTGCTGGGCT 325
QY 257 TGCTGCAGGGCCCTGGCTCGCTGCTGCTCTCTCTCTCCCTCGCCTCGCACTGCTGCTGGGCT 316
Db 326 CGCGCGCGGGCTCTTCTCTTTGGCCAGCCCGACTTCTCTACAGCGCAGCAATTGCA 385
QY 317 CGCGCGCGGGCTCTTCTCTTTGGCCAGCCCGACTTCTCTACAGCGCAGCAATTGCA 376
Db 386 AGCCCATCCGGCCAAACCTGCAGCTGTGCCACGGCATCGAATACAGAAACATGCGGCTGC 445
```

||||| 377 AGCCCATCCGGCCAACTGCAGCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGC 436
Db 446 CCAAACTGCTGGGCCACGAGACCATGAAGGAGGTGCTGGAGAGGCCGGCGCTTGGAATCC 505
||||| 437 CCAACCTGCTGGGCCACGAGACCATGAAGGAGGTGCTGGAGAGGCCGGCGCTTGGAATCC 496
Db 506 CGCTGGTCATGAAGCAGTGCCACCCGGACACCAAGAAAGTTCTGTGCTCGCTCTTCGCCCC 565
||||| 497 CGCTGGTCATGAAGCAGTGCCACCCGGACACCAAGAAAGTTCTGTGCTCGCTCTTCGCCCC 556
Db 566 CCGTCTGCCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCTCGNTGCCGTGCAGG 625
||||| 557 CCGTCTGCCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCGCTCTGCGTGCAGG 616
Db 626 TGAAGGATCGCTGGCCCCCGGTCATGTCGCCCTTC-----CCCTGGCCCCGACATGCTTG 679
||||| 617 TGAAGGACCGCTGCGCCCCCGGTCATGTCGCCCTTCGGCTTCCCTTGCCCCGACATGCTTG 676
Db 680 AGTGCAGACCGTTTCCCCAGGAAAGCTCCAAAGGTATGTGAAGCTGC AAAAATAAAATGATG 739
||||| 677 AGTGCAGACCGTTTCCCCAGGAAAGCTCCAAAGGTATGTGAAGCTGC AAAAATAAAATGATG 736
Db 740 TCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATGTGAAGCTGC AAAAATAAAATGATG 799
||||| 737 TCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATGTGAAGCTGC AAAAATAAAATGATG 796
Db 800 ATGACAACGACATAATGGAACGCTTTGTAAAAATGATTTGCACTGAAATAAAAGTGA 859
||||| 797 ATGACAACGACATAATGGAACGCTTTGTAAAAATGATTTGCACTGAAATAAAAGTGA 856
Db 860 AGGAGATAACCTACATCAACCG 881
||||| 857 AGGAGATAACCTACATCAACCG 878

RESULT 5
LOCUS AF017988 1984 bp mRNA PRI 21-SEP-1997
DEFINITION Homo sapiens secreted apoptosis related protein 3 (SARP3) mRNA,
complete cds.
ACCESSION AF017988
NID g2415418
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1984)
AUTHORS Melkonyan,H., Chang,W.C., Shapiro,J.P., Mahadevappa,M.,
Fitzpatric,P.A., Kiefer,M.C., Tomei,D.L. and Umansky,S.R.
TITLE SARPs - a new family of proteins that regulate apoptosis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1984)
AUTHORS Melkonyan,H., Prochazka,V. and Umansky,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
South, Richmond, CA 94804, USA
FEATURES
source
1. .1984
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/tissue_type="pancreas"
1. .1984
/gene="SARP3"
216. .1169
/gene="SARP3"
/note="similar to frizzled-like receptor, extracellular
protein"
/codon_start=1
/product="secreted apoptosis related protein 3"
/db_xref="PID:g2415419"
/translation="MRAAAAAGGVRTAALALLGALHWAPARCEYDYYGWQAEPLHG

RSYSKPPQCLDIPADLPLCHTVGYKRMRLPNLLEHESLAEVKQQQASSWLPLLAKRCHS
DTQVFLCSLEFAPVCLDRPIYPCRSLCEAVRAGCAPLMEAYGFPWPEMLCHKFPDND
LCIAVQFGHLPATAPPVTKICAOCEMEHSADGLMEQMCSSDFVVKMRIKEIKIENGDR
KLIGAQKKKKLLKPGPLKRKDTKRLVLHMRNGAGCPCQLDSLAGSFLVMGRKVDGQL
LLMAVYRWDKKNKMKFAVKFMFSYPCSLYIPFFYGAAEPH"
BASE COUNT 353 a 634 c 627 g 370 t
ORIGIN
Query Match 5.4%; Score 110; DB 21; Length 1984;
Best Local Similarity 73.0%; Pred.No. 8.64e-50;
Matches 246; Conservative 0; Mismatches 82; Indels 9; Gaps 3;
Db 380 CATCCCTGCCGACCTGCCGCTCTGCCACACGGTGGGTACAGCGCATGCGGTGCCCAA 439
||||| 381 CATCCGGCCAACTGCAGCTGTGCCACGGCATCGAATACCAGAACATGCGGTGCCCAA 440
Db 440 CCTGCTGGAGCACGAGAGCCCTGCCCGAAGTGAAGCAGCAGCGAGCAGCTGGCTGCCGCT 499
||||| 441 CCTGCTGGCCACGAGACCATGAAGGAGGTGCTGGAGCAGCGCGGCTTGATCCCGCT 500
Db 500 GCTGCCAAAGCGTGCCACTCGGATACGAGGTCTTCTGTGCTCGCTCTTTGCCGCCCGT 559
||||| 501 GGTATGAAGCAGTGCCACC CGGACACCAAGAAAGTTCTGTGCTCGCTCTTCGCCCGCT 560
Db 560 CTGTCTCGA-----CG-GCCCATCTACCCGTGCCGTGCTGTGCGAGGCCGTGCG 610
||||| 561 CTGCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCGCTCTCGTGCAGGTGAA 620
Db 611 CGCGGCTGCGCGCGCTCATGGAGGCCCTACGGCTTCCCTGGCCTGAGATGCTGCACTG 670
||||| 621 GGACCGCTGCGCCCCGGTCACTGTCGCCCTTCGGCTTCCCTGGCCCCGACATGCTTGAGTG 680
Db 671 CCACAAGTTCCCCCTGGACACGACCTCTGCATCGCC 707
||||| 681 CGACCGTTTCCCCCAGGACAAACGACCTTTGCATCCCC 717

RESULT 6
LOCUS AF017987 2094 bp mRNA PRI 21-SEP-1997
DEFINITION Homo sapiens secreted apoptosis related protein 2 (SARP2) mRNA,
complete cds.
ACCESSION AF017987
NID g2415416
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2094)
AUTHORS Melkonyan,H., Prochazka,V., Chang,W.C. and Umansky,S.R.
TITLE SARPs - a new family of proteins that regulate apoptosis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2094)
AUTHORS Melkonyan,H., Prochazka,V. and Umansky,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
South, Richmond, CA 94804, USA
FEATURES
source
1. .2094
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/tissue_type="heart"
1. .2094
/gene="SARP2"
303. .1247
/gene="SARP2"
/note="similar to frizzled-like receptor, extracellular
protein"
/codon_start=1
/product="secreted apoptosis related protein 2"
/db_xref="PID:g2415417"
/translation="MGIGRSEGGRRGAALGVLLALGAALLAVGSASEYDYVSFQSDIG

PYQSGRFYTKPQCVDIPADLRRLCHNVGYKKMVLPLNLEHETMAEVKQQAASSWVPLLN
KNCHAGTQVFLCSLFAPVCLDRPIYPCRWLCEAVRDSCEPVMQFFGYWPEMLKCDKF
PEGDVCIAMTPPNPTFASKPQGTTCVPPCDNELKSEAIIEHLCASEFALRMKIKEVKK
ENGDKKIIVPKKKPLKLGPIKKDKLKLVLKNGADPCPCHQDLNLSHHFLIMGRKVK
SQYLLTAIHKWDKNKFEKFMKMKMKNHECPTFQSVFK"

BASE COUNT 484 a 598 c 631 g 381 t
ORIGIN

Query Match 5.2%; Score 106; DB 21; Length 2094;
Best Local Similarity 71.8%; Pred. No. 4.01e-47;
Matches 249; Conservative 0; Mismatches 89; Indels 9; Gaps 5;

Db 452 CTTCTACACCAAGCCACCTCAGTGGCTGGACATCCCCGGGACCTCGCGGTGTGCCACAA 511
||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 351 CTTCTCTACAGCGCAGCAATTGCAAGCCCATCCCGGCCAACCTGCAGCTGTGCCACGG 410

Db 512 CGTGGGCTACAAGAAGATGGTGTGCCCCAACCTGCTGGAGCACGAGACCATGGCGGAGGT 571
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 CATGAATACCAGAACATGCGGCTGCCAACCTGCTGGGCCACGAGACCATGAAGGAGGT 470

Db 572 GAAGCAGAGGCCAGCAGCTGGGTGCCCTGCTCAACAAGAACTGCCACGCCGACCCCA 631
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 GCTGGAGCAGCGCGCGCTTGATCCCGCTGGTCATGAAGCAGTGCACCCCGGACACCA 530

Db 632 GGTCCTCTCTGCTCGCTCTTCGCGCCCGCTGCTGCCT-G--GACCG-G-C---CCATCTA 682
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 GAAGTCTCTGTGCTCGCTCTTCGCCCCCGCTGCTGCCTCGATGACCTAGACGAGACCATCA 590

Db 683 CCCGTGCTGCTGCTGCGAGCGCGTGGCGGACTCGTGGAGCGCGGTGCATGCAGTCTT 742
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 591 GCCATGCCACTCGCTCTGCGTGACGTGAAGGACCGCTGCGCCCCGCTCATGTCCGCCCTT 650

Db 743 CGGCTTCTACTGGCCCCGAGATGCTTAAGTGTGACAAAGTTTCCCGGAGG 789
|||||| | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 651 CGGCTTCCCTGGCCCCGACATGCTGAGTGGACCGCTTTCGCCCCAGG 697

RESULT 7
LOCUS BTU85945 4240 bp mRNA MAM 21-AUG-1997
DEFINITION Bos taurus frizzled-related protein Frza mRNA, complete cds.
ACCESSION U85945
NID 92337936
KEYWORDS .
SOURCE Bos taurus.
ORGANISM Bos taurus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora;
Bovidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 4240)
AUTHORS Duplaa,C. and D'Amore,P.A.
TITLE Identification and cloning of a novel secreted form of mammalian frizzled: evidence to suggest a role in the control of growth and differentiation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4240)
AUTHORS Duplaa,C. and D'Amore,P.A.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1997) INSERM U441, AV du Haut Leveque, Pessac 33600, France
FEATURES Location/Qualifiers
source 1. .4240
/organism="Bos taurus"
/db_xref="taxon:9913"
/cell_type="endothelial"
/tissue_type="heart"
198. .1124
/note="frizzled related protein"
/codon_start=1
/product="Frza"
/db_xref="PID:g2337937"
/translation="MGGRWAAAGALLAAGLLAAGSASEYDYVSFQSDIGAYQSGR
FYTKPPQCVDIPADLRRLCHNVGYKRMVLPLNLEHETMAEVKQQAASSWVPLLNKCHIG
TQVFLCSLFAPVCLDRPIYPCRWLCEAVRDSCEPVMQFFGYWPEMLKCDKFPEGDVVC

CDS

IAMTPPNATEASKPQGTTCVPPCDNELKSEAIIEHLCASEFALRMKIKEVKKENGDKK
IVPKKKKPLKLGPIKKKELKLVLYLKNGADPCPCHQDLNLSHHFLIMGRKVKSQYLLT
AIHKWDKKNKEFTFMKMKMKNHECPTFQSVFK"

sig_peptide 210. .263
misc_feature 966. .968
/note="encodes N-glycosylation site"

BASE COUNT 1015 a 1073 c 1112 g 1040 t
ORIGIN
Query Match 5.2%; Score 106; DB 15; Length 4240;
Best Local Similarity 71.4%; Pred. No. 4.01e-47;
Matches 252; Conservative 0; Mismatches 92; Indels 9; Gaps 5;

Db 329 CTTCTACACCAAGCCACCGCAGTGGCTGGACATCCCGGACACCTGCGGTGTGCCACAA 388
||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 351 CTTCTCTACAGCGCAGCAATTGCAAGCCCATCCCGGCCAACCTGCAGCTGTGCCACGG 410

Db 389 CGTGGGCTACAAGAGGATGGTGTGCCCCAACCTGCTGGAGCACGAGACCATGGCCGAGGT 448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 CATCGAATACCAGAACATGGGCTGCCAACCTGCTGGGCCACGAGACCATGAAGGAGGT 470

Db 449 GAAGCAGCAGGCCAGCAGCTGGGTGCCCTGCTCAACAAGAACTGCCACATGGCACCCCA 508
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 GCTGGAGCAGCGCGCGCTTGATCCCGCTGCTCATGAAGCAGTGCACCCCGGACACCA 530

Db 509 GGTCTTCTGTGCTCGCTCTTCGCGCCCGCTGCTGCCT-G--GACCG-G-C---CCATCTA 559
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 GAAGTTCCTGTGCTCGCTCTTCGCCCCCGCTGCTGCCTCGATGACCTAGACGAGACCATCCA 590

Db 560 CCCGTGCCGCTGCTCGAGGCCGTGCGGACTCGTGGAGCCCGTGCATGCGAGTTCCTT 619
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 591 GCCATGCCACTCGCTCTGCTGCAGGTGAAGGACCGCTGCGCCCCGCTCATGTCCGCCCTT 650

Db 620 CGGCTTCTACTGGCCCCGAGATGCTCAAGTGGCACAAAGTTTCCCGAGGGCGAGC 672
|||||| | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 651 CGGCTTCCCTGGCCCCGACATGCTTGAGTGGCACCGTTTCCCCCAGGACAACG 703

RESULT 8
LOCUS AF001900 2075 bp mRNA PRI 25-JUN-1997
DEFINITION Homo sapiens secreted frizzled-related protein mRNA, complete cds.
ACCESSION AF001900
NID 92213818
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 2075)
AUTHORS Finch,P.W., He,X., Kelley,M.J., Uren,A., Schaudies,R.P.,
Popeacu,N.C., Rudikoff,S., Aaronson,S.A., Varmus,H.E. and
Rubin,J.S.
TITLE Purification and Molecular Cloning of a Secreted, Frizzled-Related Antagonist of Wnt Action
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1997) In press
REFERENCE 2 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
TITLE -Direct Submission
JOURNAL Submitted (30-APR-1997) LCMB, DBS/NCI, 9000 Rockville Pike, Bethesda, MD 20892, USA
REFERENCE 3 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1997) LCMB, DBS/NCI, 9000 Rockville Pike, Bethesda, MD 20892, USA
REMARK Sequence update by submitter
FEATURES Location/Qualifiers
source 1. .2075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p11.1-p12"

CDS

```
/clone="HS1"
/dev_stage="embryonic"
/tissue_type="lung"
/cell_line="M426"
/cell_type="fibroblast"
303.1244
/note="FRP"
/codon_start=1
/product="secreted frizzled-related protein"
/db_xref="PID:g2213819"
/translation="MGIGRSEGGRRGALGVLLALGALLAVGSASEYDYVSFQSDIGP
YQSGRFYTKPPQCVDIPADLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNK
NCHAGTQVFLCSLFAFVCLDRPIYPCRWLCEAVRDSCEPVMOFFGYWPMLKCDKFP
EGDVCIAMTPPNATEASKPQGTVCPCDNLKSEAIIEHLCASEFALRMKIKEVKKE
NGDKKIIVPKKKPLKLGPIKKKDLKLLVLYLKNAGDCPCQDNLSHHFLIMGRKVK
QYLLTAIHKWKDKNKEFKNFMKNHECPTFQSVFK"
BASE COUNT 473 a 596 c 626 g 380 t
ORIGIN
```

Query Match 5.1%; Score 103; DB 21; Length 2075;
Best Local Similarity 73.5%; Pred. No. 3.93e-45;
Matches 161; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Db 449 CTTCTACACCAAGCCACCTCAGTGGTGGACATCCCCGGACCTGCGGCTGTGCCACAA 508
||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
QY 351 CTTCTCCTACAGCGCAGCAATTGCAAGCCCATCCCGGCCAACCTGCAGCTGTGCCACGG 410

Db 509 CGTGGGCTACAAGAAGATGGTGTGCTGCCCCACCTGCTGGAGCACGAGACCATGGCGGAGGT 568
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 CATCGAATACCAAGACATGCGGCTGCCCCAACCTGCTGGCCACGAGACCATGAAGGAGGT 470

Db 569 GAAGCAGCAGGCGCAGCAGCTGGGTGCCCCCTGTCTCAACAAGAACTGCCACGCCGGACCCA 628
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 GCTGGAGCAGGCGCGCTTGATCCCGCTGTGTCATGAAGCAGTCCACCCCGGACACCAA 530

Db 629 GGTCTTCCTCTGCTCGCTCTTCGCGCCCCGCTGCTGCTGGA 667
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 GAAGTTCTGTGCTCGCTCTTCGCCCCCGCTGCTGCTCGA 569

RESULT 9
LOCUS AF056087 4469 bp mRNA PRI 08-APR-1998
DEFINITION Homo sapiens secreted frizzled related protein mRNA, complete cds.
ACCESSION AF056087
NID g3033550
KEYWORDS .
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4469)
AUTHORS Zhou,Z. and Wang,J.
TITLE Upregulation of human secreted Frizzled homologue in apoptosis and its down regulation in breast tumors

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4469)
AUTHORS Zhou,Z. and Wang,J.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1998) Radiohemmet Research Laboratory, CCK, Karolinska Institute, Solnavagen, Stockholm S-171 76, Sweden
FEATURES Location/Qualifiers
source 1. .4469
/organism="Homo sapiens"
/db_xref="taxon:9606"
303.1244

CDS
/function="putative Wnt antagonist"
/note="FRP; frizzled homolog; cell cycle and apoptosis related; down regulated in breast tumoregenesis"
/codon_start=1
/product="secreted frizzled related protein"
/db_xref="PID:g3033551"
/translation="MGIGRSEGGRRGALGVLLALGALLAVGSASEYDYVSFQSDIGP
YQSGRFYTKPPQCVDIPADLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLLNK

NCHAGTQVFLCSLFAFVCLDRPIYPCRWLCEAVRDSCEPVMOFFGYWPMLKCDKFP
EGDVCIAMTPPNATEASKPQGTVCPCDNLKSEAIIEHLCASEFALRMKIKEVKKE
NGDKKIIVPKKKPLKLGPIKKKDLKLLVLYLKNAGDCPCQDNLSHHFLIMGRKVK
QYLLTAIHKWKDKNKEFKNFMKNHECPTFQSVFK"
BASE COUNT 1121 a 1097 c 1120 g 1131 t
ORIGIN

Query Match 5.1%; Score 103; DB 21; Length 4469;
Best Local Similarity 73.5%; Pred. No. 3.93e-45;
Matches 161; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Db 449 CTTCTACACCAAGCCACCTCAGTGGTGGACATCCCCGGACCTGCGGCTGTGCCACAA 508
||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
QY 351 CTTCTCCTACAGCGCAGCAATTGCAAGCCCATCCCGGCCAACCTGCAGCTGTGCCACGG 410

Db 509 CGTGGGCTACAAGAAGATGGTGTGCTGCCCAACCTGCTGGAGCACGAGACCATGGCGGAGGT 568
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 411 CATCGAATACCAAGACATGCGGCTGCCCAACCTGCTGGGCCACGAGACCATGAAGGAGGT 470

Db 569 GAAGCAGCAGGCGCAGCAGCTGGGTGCCCCCTGTCTCAACAAGAACTGCCACGCCGGACCCA 628
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 GCTGGAGCAGGCGCGCTTGATCCCGCTGATCCCGTGCATGAAGCAGTGCCACCCGGACACCAA 530

Db 629 GGTCTTCCTCTGCTCGCTCTTCGCGCCCCGCTGCTGCTGGA 667
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 GAAGTTCTGTGCTCGCTCTTCGCCCCCGCTGCTGCTCGA 569

RESULT 10
LOCUS MMU88566 2659 bp mRNA ROD 22-APR-1997
DEFINITION Mus musculus secreted frizzled related protein sFRP-1 (Sfrpl) mRNA, complete cds.
ACCESSION U88566
NID g1946340
KEYWORDS .
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 2659)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)
MEDLINE 97250455
REFERENCE 2 (bases 1 to 2659)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns Hopkins University School of Medicine, 725 North Wolfe Street, PCTB 805, Baltimore, MD 21205
FEATURES Location/Qualifiers
source 1. .2659
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="8"
/map="between Plat and Fgfr1"
257.1201
/gene="Sfrpl"
257.1201
/gene="Sfrpl"
/codon_start=1
/product="secreted frizzled related protein sFRP-1"
/db_xref="PID:g1946341"
/translation="MGVGRSARGRGAASGVLLALAAALGAASASEYDYVSFQSDIG
SYQSGRFYTKPPQCVDIPVDLRLCHNVGYKKMVLPNLLEHETMAEVKQQASSWVPLIN
KNCHMGTOVFLCSLFAFVCLDRPIYPCRWLCEAVRDSCEPVMOFFGYWPMLKCDKF
PEGDVCIAMTPPNTEASKPQGTTCPCDNLKSEAIIEHLCASEFALRMKIKEVKK
ENGDKKIIVPKKKPLKLGPIKKKELKALVLFKNGADCPCHQDNLSHNELIMGRKVK
SOYLLTAIHKWKDKNKEFKNFMKRMKNHECPTFQSVFK"


```
/db_xref="taxon:9031"
/clone="5.2"
23. .946
/note="member of cysteine-rich frizzled family of
proteins"
/codon_start=1
/product="crescent"
/db_xref="pid:g2226372"
/translation="MRAASTEKASRRRGAPWLPGLLLRLLWGAGGRASYLRRSSSC
TAIPRSMALCYDIGYSEMRIPNLLEHETMPEVIQQSSSWLPLLARECHDPDAIFLCSL
FAPICDLRLIYPCSLCEAVKRCAPVMACYGYPWPPEILNCNKFPPADHELICIAAVSTD
ESSSSRRMPRASCCKDELEEAAREILDNLCAANDFTVKIRILRKNNTTTVSDFDLDP
SRVEVLKPGPLRLTEIPGRLOQWLIDIDATCAHNIMRGTHAGVFVISGEVRS DKVVVNK
AYAWQKKNRNLHOAVRRWKHRCPEQAGRKV"
BASE COUNT      517 a      633 c      547 g      378 t
ORIGIN

Query Match      4.1%; Score 83; DB 16; Length 2075;
Best Local Similarity 71.1%; Pred. No. 4.64e-32;
Matches 140; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Db 152 TGCACCGCCATCCCGCGCAGCATGGCCCTGTGCTATGACATCGGTTACTCGGAGATGAGG 211
||||| ||||| || || ||||| || ||||| || ||||| || ||||| ||
QY 373 TGC AAGCCCATCCCGGCCAACCTGCGAGCTGTGCCACGGCATCGAATACCAGAACATGCCG 432

Db 212 ATCCCCAACCTGCTGGAGCACGAGACCATGCCGAGGTGATCCAGCAGTCCTCCAGCTGG 271
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 433 CTGCCAACCTGCTGGGCCACGAGACCATGAAGGAGGTGCTGGAGCAGGCCGCGCTTG 492

Db 272 CTGCCCTTGCTGGCCAGGAGTGCCATCCGACGCCAGGATTTCCTCTCTGCTCCCTCTTC 331
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 ATCCCGCTGGTCATGAAGCAGTGCCACCGGACACCAAGAAGTCTCTGTGCTCGCTCTTC 552

Db 332 GCGCCCATCTGCCTTGA 348
||| ||| ||||| |||
QY 553 GCGCCCGTCTGCCTCGA 569

RESULT 14
LOCUS HSU82169 2184 bp mRNA PRI 26-MAR-1997
DEFINITION Human frizzled homolog (FZD3) mRNA, complete cds.
ACCESSION U82169
NID g1906597
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2184)
AUTHORS Wang,Y.K., Samos,C.H., Peoples,R., Perez-Jurado,L.A., Nusse,R. and
Francke,U.
TITLE A novel human homologue of the Drosophila frizzled wnt receptor
gene binds wingless protein and is in the Williams syndrome
deletion at 7q11.23
JOURNAL Hum. Mol. Genet. 6 (3), 465-472 (1997)
MEDLINE 9727293
REFERENCE 2 (bases 1 to 2184)
AUTHORS Wang,Y.-K., Peoples,R., Perez-Jurado,L.A. and Francke,U.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1996) Howard Hughes Medical Institute, Stanford
Medical Center, Beckman Center B201, Stanford, CA 94305, USA
FEATURES
source
1. .2184
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q11.23"
25. .1801
/gene="FZD3"
26. .1801
/gene="FZD3"
/note="similar to Drosophila frizzled; seven transmembrane
receptor"

/db_xref="taxon:9031"
/clone="5.2"
23. .946
/note="member of cysteine-rich frizzled family of
proteins"
/codon_start=1
/product="crescent"
/db_xref="pid:g2226372"
/translation="MRAASTEKASRRRGAPWLPGLLLRLLWGAGGRASYLRRSSSC
TAIPRSMALCYDIGYSEMRIPNLLEHETMPEVIQQSSSWLPLLARECHDPDAIFLCSL
FAPICDLRLIYPCSLCEAVKRCAPVMACYGYPWPPEILNCNKFPPADHELICIAAVSTD
ESSSSRRMPRASCCKDELEEAAREILDNLCAANDFTVKIRILRKNNTTTVSDFDLDP
SRVEVLKPGPLRLTEIPGRLOQWLIDIDATCAHNIMRGTHAGVFVISGEVRS DKVVVNK
AYAWQKKNRNLHOAVRRWKHRCPEQAGRKV"
BASE COUNT      517 a      633 c      547 g      378 t
ORIGIN

Query Match      4.1%; Score 83; DB 16; Length 2075;
Best Local Similarity 71.1%; Pred. No. 4.64e-32;
Matches 140; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Db 152 TGCACCGCCATCCCGCGCAGCATGGCCCTGTGCTATGACATCGGTTACTCGGAGATGAGG 211
||||| ||||| || || ||||| || ||||| || ||||| || ||||| ||
QY 373 TGC AAGCCCATCCCGGCCAACCTGCGAGCTGTGCCACGGCATCGAATACCAGAACATGCCG 432

Db 212 ATCCCCAACCTGCTGGAGCACGAGACCATGCCGAGGTGATCCAGCAGTCCTCCAGCTGG 271
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 433 CTGCCAACCTGCTGGGCCACGAGACCATGAAGGAGGTGCTGGAGCAGGCCGCGCTTG 492

Db 272 CTGCCCTTGCTGGCCAGGAGTGCCATCCGACGCCAGGATTTCCTCTCTGCTCCCTCTTC 331
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 ATCCCGCTGGTCATGAAGCAGTGCCACCGGACACCAAGAAGTCTCTGTGCTCGCTCTTC 552

Db 332 GCGCCCATCTGCCTTGA 348
||| ||| ||||| |||
QY 553 GCGCCCGTCTGCCTCGA 569

RESULT 15
LOCUS AC005049 106049 bp DNA HTG 12-JUN-1998
DEFINITION Homo sapiens clone RG023I15; HTGS phase 1, 1 unordered pieces.
ACCESSION AC005049
NID g3212942
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106049)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 106049)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
* This sequence is unfinished. It consists of 1 contigs for
* which the order is not known; their order in this record is
* arbitrary. In some cases, the exact lengths of the gaps
* between the contigs are also unknown; these gaps are presented
```

```
/codon_start=1
/product="frizzled homolog"
/db_xref="pid:g1906598"
/translation="MAVAPLRGALLLWQLLAAGGALEIGREFDPERGRGAAPCAQAVEI
PMCRGIGYNLTRMPNLIGHTSQGEAAAEAEFAPLVQYGCCHSLRFFLCSLYAPMCTD
QVSTPIACRPMCEQARLCAPIMEQNFNGWPDSDLCARLPTRNDPHALCMEAPENAT
AGPAEPHKGLGMLVAPRPAPPPDGLPGAGGSGTCENPEKFQYVEKSRSCAPRCGPG
VEVFWRRDKDFALVMAVWSALCFFSTAFTVLTFLLEPHRFQYPERPIIFLSMCYNV
YSLAEFLIRAVAGAQSVACDQEGALYVIOEGLENTGCTLVFLLLYFEGMASSLWMVVL
TLTWFLAAGKKWGHEAIEAHGSYFHMAAWGLPALKTIVILTLRKVAGDELTLGCVVAS
TDAALTTGFVLVPLSGYLVLSGSELLTGTFVALFHIRKIMKTGGTNTKLEKLMVKIGV
FSILYTPATCVIVYVERLNMDFWRLRATEQPCAAAAGPGRRDGCSLPGGSVPTVA
VFMLKIFMSLVVGTSGVWVWSSKTFQTWQSLCYRKIAAGRARAKACRAPGSGYGRGTH
CHYKAPTIVLHMTKIDPSLENPTHL"
BASE COUNT      329 a      737 c      729 g      389 t
ORIGIN

Query Match      3.0%; Score 60; DB 21; Length 2184;
Best Local Similarity 59.8%; Pred. No. 1.15e-17;
Matches 183; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

Db 162 TGTGCGCGGCATCGGCTACAACTGACCCGCGCATGCCCAACCTGCTGGGCCACACGTCGC 221
||||| ||||| ||| || || ||||| ||||| ||||| ||||| ||||| ||
QY 401 TGTGCCACGGCATCGAATACCAGAACATCGGGCTGCCCAACCTGCTGGGCCACGAGACCA 460

Db 222 AGGCGAGGCGGCTGCCGAGCTAGCGGAGTTCGCCCGCTGGTGAGTACGGCTGCCACA 281
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 461 TGAAGGAGGTGCTGGAGCAGGCCGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACC 520

Db 282 GCCACCTGCGCTTCTTCTGCTGCTCGCTCTACGCGCCCATGTGCACCGCAGGCTCTCGA 341
|| || || || || || || || || || || || || || || || || || ||
QY 521 CGGACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGCTCGCTCGATGACCTAGACG 580

Db 342 CGCCCATTCGCCCTGCCCGCCCATGTGCGAGCAGGCGCGCTGGCTGCGCGCCCATCA 401
|| || || || || || || || || || || || || || || || || || ||
QY 581 AGACCATCCAGCCATGCCACTCGCTCTGCGTGCAAGGACCGCTGCGCCCGGTCA 640

Db 402 TGGAGCAGTTCAACTTCGGCTGGCGGAGTTCGCTCGACTGCGCGCGGCTGCCACGCGCA 461
|| || || || || || || || || || || || || || || || || || ||
QY 641 TGTCGCGCTTCGGCTTCCCTGGCCCGACATGTTGAGTGCAGACCGTTTCCCCCAGGACA 700

Db 462 ACGACC 467
|||||
QY 701 ACGACC 706

RESULT 15
LOCUS AC005049 106049 bp DNA HTG 12-JUN-1998
DEFINITION Homo sapiens clone RG023I15; HTGS phase 1, 1 unordered pieces.
ACCESSION AC005049
NID g3212942
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106049)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 106049)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
* This sequence is unfinished. It consists of 1 contigs for
* which the order is not known; their order in this record is
* arbitrary. In some cases, the exact lengths of the gaps
* between the contigs are also unknown; these gaps are presented
```



```
/product="transmembrane receptor"
/db_xref="pid:g1151252"
/translation="MARPDPSAPPSLLLLLAQLVGRAAAASKAPVCOEITVPMCRGI
GYNLTHMPNQFNHDTQDEAGLEVHQFPLVEIQCSPLDFLFFCTMYTPICLPDYHKPL
PPCRSVCKERAKAGCSPLMRQYGFQWPERMSCDRLPVLGRDAEVLMDYNRSEATTAPP
RPFPAKPTLPGPAPASGCECPAGPFVCKREPFVILKESHPLXNKVRTGQVPNC
AVPCYQPSFADERTFATFWIGLSVLCFISTTIVATFLIDMTFRPERPIIFLSA
CYLCSVLGFLVRLVVGHASVACRSREHNHIVETGTPALCTIVFLLVYFFGMASSIWWV
ILSLTWFLAAMKWGNEAIAAGYQYFHLAHLIPSVKSITALSSVDGDPVAGICYV
GNQNLNSLRRFVLGPLVLYLLVGLFLLAGFVSLFRIRSVIKOGGTDKLEKLMIRI
GIFTLTYTPASIVVACYLEQHYRESWEALTCACPGHDTGQPRAKPEYVWVLMKYF
MCLVVGITSGVMTWSGKTVESWRRFTSRCCRPRRGHKSGGAMAAGDYPEASALTGR
TGPPGPAATYHKQVSLSHV"
BASE COUNT 356 a 803 c 736 g 439 t
ORIGIN
```

```
Query Match 2.9%; Score 58; DB 20; Length 2334;
Best Local Similarity 59.9%; Pred. No. 1.84e-16;
Matches 175; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Db 439 TGTGCGCGGCATCGGCTACACCTGACGCACATGCCCCACCAAGTTCAACCACGAGACGC 498
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 401 TGTGCCACGGCATCGAATACCAACAACATGCGGCTGCCCAACCTGCTGGCCACGAGACCA 460
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 AGGACGAGCGGGCTGGAGGTGCACCAAGTTCTGGCCGCTGGTGGAGATCCAATGCTCGC 558
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 461 TGAAGGAGGTGCTGGAGCAGCGCCGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACC 520
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 CGGACCTGCGCTTCTTCCTATGCACTATGTACACGCCCACTGTCTGCCGACTACCACA 618
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 521 CGGACACCAAGAAGTTCTTGCTGCTGCTCTTCGCGCCCGCTGCTGCTGATGACCTAGACG 580
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 619 AGCGCTGCGCGCCCTGCGCTGCGGTGTGCGAGCGGCGCAAGCGCGGCTGCTGCGCGGTGA 678
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 581 AGACCATCCAGCCATGCCACTCGCTGTGCGTGCAGGTGAAGGACCGCTGCGCCCGGTCA 640
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 679 TCGCGAGTACGGCTTCGCCTGCGCCGAGCGGATGAGTGGACCGCCTCCC 730
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 641 TGTCCGCTTCGGCTTCGCCCTGCGCCCGACATGCTTGAGTGGACCGTTTCCC 692
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 18
LOCUS AF031830 2328 bp mRNA VRT 02-DEC-1997
DEFINITION Gallus gallus 7-transmembrane protein frizzled-1 (Fz-1) mRNA,
nuclear gene encoding mitochondrial protein, complete cds.
ACCESSION AF031830
NID 92655273
KEYWORDS chicken.
SOURCE Gallus gallus
ORGANISM Eukaryotae; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2328)
AUTHORS Kengaku,M., Twombly,V. and Tabin,C.
TITLE Expression of Wnt and Frizzled genes during chick limb bud
development
JOURNAL Cold Spring Harb. Symp. Quant. Biol. (1997) In press
REFERENCE 2 (bases 1 to 2328)
AUTHORS Kengaku,M. and Tabin,C.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1997) Genetics, Harvard Medical School, 200
Longwood Ave., Boston, MA 02115
FEATURES
source
1. .2328
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="limb bud"
1. .2328
/gene="Fz-1"
212. .1990
/gene="Fz-1"
/note="putative Wnt receptor; similar to rat frizzled-1
protein"
```

```
/codon_start=1
/product="7-transmembrane protein frizzled-1"
/db_xref="pid:g2655274"
/translation="MAERRGPAGGSGEVGGRRAGGDRCPRRPPALPLLWAAAL
PAGQPAQAQPAALSERGISIPDHGYCQIPISPLCTDIAYNQIMPNNLLHTNQEDAGL
EVHQFYPLVKQCSAELKFFLCMYAPVCTVLEQALPPCRSLCERARQGCCEALMNKFG
FQWPDTLRCEKFPVHGAGELCVQNASERGTPTPALRPESWTSNPHRGGGAGSGPGE
ARGFSCPRALKVPVSYLNYRFLGEKDCGAPCEPGRLYGLMYFGPEELRFSRTWIGWS
VLCCASTLFTVLTLYLVDMKRFYSYPERPIIFLSCYTAVAVAYIAGFLLEERVVCNERF
AEDGSTVAOQTKREGCTILFMMLYFFGMASSIWWVILSLTWFLAAGMKWGHEAIEAN
SOYFHLLAAWAVPAIKTITILALQVQDGVLSGVCFVGINNVDALRGFVLAFLFVYLF
GTSFLLAGFVSLFRIRTIMKHDGKTKELEKLMVRIGIFSPLYTVPATIVACYFYEQ
AFREQWERSWVTQSKSYAIPCNNHSSHHPMPSPDFTVFMIKYMLTILVIGITSGFWI
WSGKTLNSWRKFYTRLTNSKQGETTV"
BASE COUNT 394 a 732 c 737 g 463 t
ORIGIN
```

```
Query Match 2.8%; Score 57; DB 16; Length 2328;
Best Local Similarity 62.2%; Pred. No. 7.29e-16;
Matches 183; Conservative 0; Mismatches 108; Indels 3; Gaps 2;

Db 439 GCTGTGCACAGACATCGCTTACAACCAGACCATCATGCCCAACCTGCTGGGCCACACCAA 498
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 399 GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGAC 458
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 CCAGGAGGACGCGGGCTGGAGGTGCACCAAGTTCTACCCGCTGGTGAAGGTGCAGTGCTC 558
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 459 CATGAAGGAGGTGCTGGAGCAGCGCGCGCTTGGATCCCGTGGTCATGAAGCAGTGCCA 518
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 GGCCGAACCTGAAGTTCTTCTGTGTCTCCATGTACGCCCGCGGTGTGCACCG-TG--CTGGA 615
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 519 CCGGACACCAAGAAGTTCTGTGTCTGCTGCTTTCGCCCGCTGCTGCTCGATGACCTAGA 578
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 616 GCAGGCCCTGCCGCCCTGCCGCTCTCTTTGCGAGCGCGCGCCAGGGCTGCGAGGCCCT 675
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 579 CGAGACCATCCAGCCATGCCACTGCTGCTGCGTGCAGTGAAGGACCGCTGCGCCCGCGT 638
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 676 CATGAACAAGTTGCGGCTTCCAGTGGCCCGACACGCTGCGCTGCGAGAAAGTTCCC 729
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 639 CATGTCGCGCTTCGGCTTCCCTTCCCTGGCCCGACATGCTTGAGTGGACCGCTTCCC 692
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 19
LOCUS MMU43321 2421 bp DNA ROD 24-FEB-1996
DEFINITION Mus musculus putative transmembrane receptor (frizzled 8) gene,
complete cds.
ACCESSION U43321
NID 91151259
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 2421)
AUTHORS Wang,Y., Macke,J.P., Abella,B.S., Andreasson,K., Worley,P.,
Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE A large family of putative transmembrane receptors homologous to
the product of the Drosophila tissue polarity gene frizzled
JOURNAL J. Biol. Chem. 271 (8), 4468-4476 (1996)
MEDLINE 96224032
REFERENCE 2 (bases 1 to 2421)
AUTHORS Abella,B., Wang,Y., Macke,J.P. and Nathans,J.
TITLE Direct Submission
JOURNAL Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and
Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street,
Baltimore, MD 21205
FEATURES
source
1. .2421
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="18"
/map="between Tpl2 and Cdh2"
```



```

/ gene="szl"
20. .865
/ gene="szl"
/ codon_start=1
/ product="secreted xwnt8 inhibitor sizzled"
/ db_xref="PID:g3089551"
/ translation="MSGVFLLLCASM LACTA AFDIGLSTKCVTIPTEMAMCNDVG YSE
MRLPNLMGHTNMAEVVPKSAEWNLIQTGCHPYARTFLCSLFAPVLDTFIQPCRS MC
VLRDSCAPVLACHGHSWPESLDCDRFPAGEDMCLDTLSKEYQVSYKELPKPSCQGCP
LIEEFESHKTVLEAFCDNNFAVKVRLAKKSASGLYEYETEGVEFIKQGLLLPYDTR
TMIEQWLLINENCAQKLIRTPTVVIAGEIHHGKVKNRIFHWQKKDSQLTLATRRW
RHHKC"
BASE COUNT      357 a      256 c      253 g      315 t
ORIGIN
Query Match      2.4%; Score 48; DB 16; Length 1181;
Best Local Similarity 61.8%; Pred. No. 1.32e-10;
Matches 126; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Db 91 CACCAAGTGGCTACCAATCCACAGAGATGGCCATGTGCAATGACGTGCGTACTCGGA 150
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 366 CAGCAATTGCAAGCCCATCCCGGCCAACCTGCAGCTGTGCCACGGCATCGAATACGAGAA 425
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 GATCGGGCTGCCAAACCTGATGGGACACACTAACATGSCAGAAAGTCGTGCCCAAGTCAGC 210
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 CATCGGGCTGCCCAACCTGTGGCCACGAGACCATGAAGGAGGTGCTGGAGCAGCCGG 485
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 AGAGTGGCAGAACCTCTACAAACCGGCTGCCACCCCTATGCCAGGACCTTCCTGTGCTC 270
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 486 CGCTTGGATCCCGCTGTCATGAAGCAGTGCCACCCCGGACACCAAGAAGTTCCTGTGCTC 545
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 CCTGTTCCGCCCGGCTGCCTGGA 294
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 546 GCTCTTGGCCCCCGCTGCTGCTCGA 569

RESULT 22
LOCUS AF017177 1851 bp mRNA VRT 11-AUG-1998
DEFINITION Xenopus laevis frizzled 8 protein (Xfz8) mRNA, complete cds.
ACCESSION AF017177
NID 9323437
KEYWORDS .
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia;
Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 1851)
AUTHORS Itoh,K., Jacob,J. and Sokol,S.Y.
TITLE A role for Xenopus Frizzled 8 in dorsal development
JOURNAL Mech. Dev. 74, 145-157 (1998)
REFERENCE 2 (bases 1 to 1851)
AUTHORS Itoh,K., Jacob,J. and Sokol,S.Y.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-1997) Microbiology and Molecular Genetics,
Harvard Medical School and Beth Israel Deaconess Medical Center,
330 Brookline Ave., Boston, MA 02215, USA
FEATURES
source Location/Qualifiers
1. .1851
/organism="Xenopus laevis"
/db_xref="taxon:8355"
1. .1851
/ gene="Xfz8"
1. .1746
/ gene="Xfz8"
/ codon_start=1
/ product="frizzled 8 protein"
/db_xref="PID:g323438"
translation="MESLSLSLLLVSWLQGSQCAAAKELSCOBITVPLCKDIGNNT
YMPQFNHDTQDEAGMEVHQFWPLVIHCSPLDKFFLCSMYTICLEDYKKPLPPCRS
VCERARAGCAPLMROYGFAWDRMRCDRLPEQGNPDTLCDMDYNNRTEQTAAAPSHPEP
PKPPARSPKGRTRVEPPRSRATGCESGCQCRAPMVQVSNRHPLYNVRTGQIPN
CAMPCHNPFFSPEERTFTEFWIGLWSVLCFASTFATVSTFLIDMERFKYPERPIIFLS
ACYLLVSTGYLIRLIAGHEKVACSRGELDLLEHIIHYETTGPALCTLVFLLIYFFGMAS

```

```

SIWWIILSTWFLAAGMKWGNEAIAGYSQYFHLLAOWLPSIKSIASIAVLALSSVDGDPVA
GICFVGNONLDNLRGFLVLAPLYIYLFIGSMFLLAGFVSLFRIRSVIKOGTKTKLEK
LMIRIGIFSVLYTPATIVVACFFYEQHNRRQGWVEVAHNCNSCQPEMAQPHRPDYAVFM
LKYFMCILVVGITSGVWISWGKTLSESWRAFCTRCCWGSKATGGSYSDVSTGLTWRSGT
GSSVSCPKQMPLSQV"
BASE COUNT      374 a      547 c      502 g      428 t
ORIGIN
Query Match      2.4%; Score 49; DB 16; Length 1851;
Best Local Similarity 59.1%; Pred. No. 3.55e-11;
Matches 159; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Db 103 CTGTGCAAAGACATCGGCTACAACTACACGTACATGCCCAACCAAGTTCAAACCACGACAG 162
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 CTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGACC 459
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 CAGGACGAGGCCGGGATGGAGGTGCATCAGTTCTGGCGCTGGTCGTTATTCACTGCTCC 222
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 ATGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCAC 519
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 CCGGACCTCAAGTCTTCTGTGTCAGCATGTACACCCCATCTGCCTGGAGGACTACAAG 282
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 CCGGACACCAAGAAGTCTGTGTCGCTCTTCGCCCGCGTCTGCCTCGATGACCTAGAC 579
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 AAGCCGCTGCCCCGTGTCGGAGCGTGTGTAGCGGCGCGGGCTGGTTGTGCCCTCTC 342
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 GAGACCATCCAGCCATGCCACTCGCTCTCGTGCAGGTGAAGGACCGCTGCGCCCCGGTC 639
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 343 ATGCGTCAGTACGGCTTCGCTTGGCCGGA 371
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 ATGTCGGCTTCGGCTTCCCTGGCCCCGA 668

RESULT 23
LOCUS AF012089 10772 bp DNA INV 05-AUG-1997
DEFINITION Drosophila melanogaster cysteine proteinase-1 (Cp1) gene, complete
cds, and phenylalanyl tRNA synthetase gene, partial cds.
ACCESSION AF012089
NID 92305220
KEYWORDS .
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 4546 to 4553)
AUTHORS Gray,Y.H., Tanaka,M.M. and Sved,J.A.
TITLE P-element-induced recombination in Drosophila melanogaster: hybrid
element insertion
JOURNAL Genetics 144 (4), 1601-1610 (1996)
MEDLINE 97132596
REFERENCE 2 (bases 1 to 10772)
AUTHORS Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE Structure of the cysteine proteinase (Cp1) gene of Drosophila
melanogaster and associated mutational effects
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 10772)
AUTHORS Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1997) School of Biological Sciences, University
of Sydney, Biology A12, Sydney University, NSW 2006, Australia
FEATURES
source Location/Qualifiers
1. .10772
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
join(872. .1000,2310. .2426,6476. .6690,6751. .7707)
/ gene="Cp1"
/ product="cysteine protease"
872. .7707
/ gene="Cp1"
872. .1000
/ gene="Cp1"
/ number=1

```


TDGWTNVTPKILSYVGANKHLQTDHPLSIIRQIRIVNYFYGAYRNQRGNPLFSVYDQMN
PVVTVQQNFEDNLLIPADHVSROKSDCYINQQHLLRAHTTAHQVELISGGLDNFLVVG
EYRRDEIDSTHYPVFHOADAVRLVTKDLFERNPGLELFEETWSGTLDADPKLILPH
SSWTKPNSPATRRAVKLMEHEMKHVLVGLTKDLFGPRIKYRWVDTYPFPTQPSWELEI
YFKDNWLEVLGGGIMRHEILQORSGVHQSIGYAFGVGLERLAMVLFDPDIRLFSWNS
GFLSQFSEKDLHNLPKYKPISHYPQCTNLSFWLPQDIEVDAGFSPNDFYDLVRSVAG
DMVEQISLVDKFKHPKTGKSSVCFRIVYRHMERLTLTQAEVNEIHKQIASASVDSFNVQ
IX"

BASE COUNT 2929 a 2357 c 2282 g 3046 t 158 others
ORIGIN

Query Match 2.3%; Score 46; DB 14; Length 10772;
Best Local Similarity 21.3%; Pred. No. 1.79e-09;
Matches 32; Conservative 71; Mismatches 45; Indels 2; Gaps 2;

Db 1689 AAKRWYAWTTTMMKMWTTWKWAMKTYRTWWMKMYWTSRTTTTSAMWMYTWSW 1748

QY 1790 AAAGATATAATTTTATATATTGACTTGAGTACTTTAAGCCTTGTTTAAACATTTCTTAC 1849

Db 1749 TKYWWAYAW-MKMWWTRTWARMASWARWKWKTSAAYSAWRKMW-KWWAYRAMKKTW 1806

QY 1850 TTAACTTTTCAAATTAACCCATTTGAGCTTACCTGTAATACATACATAGTATTACCTT 1909

Db 1807 MWAANKWRKAAWTWRWYMTTAAAMRAA 1836

QY 1910 TAAAAGTTGTAAAAATATTGCTTTAACCAA 1939

RESULT 25
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.

ACCESSION AC005369

NID 93367505

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 74371)

AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.

Sequencing of human chromosome 5

Unpublished

REFERENCE 2 (bases 1 to 74371)

AUTHORS Ricke,D.O.

TITLE Large Scale Sequence Analysis and Annotation with the Sequence

Comparison Analysis (SCAN) System

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 74371)

AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.

TITLE Direct Submission

JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.

COMMENT Sequence submitted by:

DOE Joint Genome Institute.

FEATURES Location/Qualifiers

source 1..74371

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="5q"

/clone="119j3"

/chromosome="5"

/note="LBNL H175"

repeat_region 893..1030

/rpt_family="Alu"

repeat_region 2295..2438

/rpt_family="Alu"

repeat_region 2818..2859

misc_feature /note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
join(3246..3410,3721..3828)
/standard_name="RLF"
/note="65% & 69% protein identity GenPept:U22377"
3431..3724
/rpt_family="Alu"
3707..3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
4366..4661
/rpt_family="Alu"
5327..5602
/rpt_family="Alu"
6586..6956
/rpt_family="L1"
6647..6684
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
7113..7373
/rpt_family="Alu"
complement(7830..8185)
/standard_name="possible repeat"
8258..8503
/rpt_family="Alu"
9070..9387
/rpt_family="Alu"
complement(9740..9845)
/rpt_family="MER42"
complement(10440..11015)
/rpt_family="Alu"
11950..12250
/rpt_family="Alu"
12057..12085
/note="(A)29"
/rpt_type=tandem
/rpt_unit=A
12365..12645
/rpt_family="Alu"
13727..13750
/note="(AC)12"
/rpt_type=tandem
/rpt_unit=AC
13783..14024
/rpt_family="L1"
14175..14470
/rpt_family="Alu"
complement(14906..15259)
/standard_name="possible repeat"
15300..15613
/rpt_family="Alu"
16671..16690
/note="(T)20"
/rpt_type=tandem
/rpt_unit=T
complement(16675..16977)
/rpt_family="Alu"
complement(16993..17085)
/rpt_family="MER42"
complement(17678..18276)
/rpt_family="Alu"
19305..19583
/rpt_family="Alu"
19914..19945
/note="(GTTT)8"
/rpt_type=tandem
/rpt_unit=GTTT
complement(19943..20222)
/rpt_family="Alu"
21202..21496
repeat_region

PVCTVLEQAIPPCRSICERAFQGC EALMNKFGFQWPERLCREHFFPRHGAEQICVGNQH
SEDTGPALLTTAPPSGLPGAGTGGPGGGGAPPRYATLEHPFHCPRVLKVPSYLSY
KFLGERDCAAPCEPARDGSMFFSHHTFARLWILTSVSLCASTFTVTTSLVAMQ
RFRYPERPIFLSGPYNMSVAYIAGFVLQERVVCNERFSDCYVAVQGTKEGCTI
LFMMLYFFSMASSIWWILSLTWFLAAGMKWGHA AEANSQYFHLAAWAPPAVKTTI
LAMQIDGDLLSGVCFVGLNRDPLRGFVLAPLFVLYFIGHTSLLAGFVSLFRITIM
KHDGKTTEPLERLMVRIGFVSVLTVPATIV IACYFYEQAQREHWERSWVSQHCKSLA
IPCBAHYTPRTSPDFTVYMIKYLMTLIVGITSGFWIWSGKTLHWSRKKFYTRLTNSRHG
ETTV *

BASE COUNT	299 a	662 c	566 g	385 t
------------	-------	-------	-------	-------

BASE C
ORIGIN

Query Match	2.2%	Score 45;	DB 22;	Length 1912;
Best Local Similarity	60.2%	Pred. No. 6.52e-09;		
Matches 177;	Conservative	0;	Mismatches 114;	Indels 3;
			Gaps 2;	

Dbb	228	GCTGTGCACGGACATCGCCTACAAACCAGACCATCATGCCAAACCTTCTTGGGCACAGAA	287
QY	399	GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCAAACCTGCTGGGCCACGAGAC	458
Dbb	288	CCAAGAGGACGCGGGCCTGGAGGTGCATCAATTCACCGCTGGTGAAGGTGCAGTGCTC	347
QY	459	CATGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCGCGCTGGTCATGAAGCAGTGCCA	518
Dbb	348	GCCCGAGCTGCGCTTCTTCGTGTGCTCCATGTACGCTCGCGTGTGCA-CGGTG--CTGGA	404
QY	519	CCCGGACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCGCGTCTGCGCTCGATGACCTAGA	578
Dbb	405	GCAGGCCATCCCGCGCTGCCGCTCCATCTGCGAACGCGCGCGCCAAAGGCTGCGAGGCGCT	464
QY	579	CGAGACCATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGT	638
Dbb	465	CATGAACAAGTTCGGCTTCCAGTGGCCCCGAGCGGCTTCGGCTGCGAGCATTTCCC	518
QY	639	CATGTCGGCTTCCCGCTTCCCGTGGCCCCGACATGCTTGAGTGCAGCCGTTTCCC	692

```

RESULT 27
LOCUS AF031831 1727 bp mRNA VRT 02-DEC-1997
DEFINITION Gallus gallus 7-transmembrane protein frizzled-7 (Fz-7) mRNA,
nuclear gene encoding mitochondrial protein, complete cds.
ACCESSION AF031831
NID g2655275
KEYWORDS .
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryotae; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1727)
AUTHORS Kengaku,M., Twombly,V. and Tabin,C.
TITLE Expression of Wnt and Frizzled genes during chick limb bud
development
JOURNAL Cold Spring Harb. Symp. Quant. Biol. (1997) In press
REFERENCE 2 (bases 1 to 1727)
AUTHORS Kengaku,M. and Tabin,C.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1997) Genetics, Harvard Medical School, 200
Longwood Ave., Boston, MA 02115
FEATURES
source Location/Qualifiers
1..1727
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="limb bud"
1..1727
/gene="Fz-7"
15..1718
/gene="Fz-7"
/note="putative Wnt receptor; similar to mouse frizzled-7
protein"
/codon_start=1
/product="7-transmembrane protein frizzled-7"
/db_xref="PID:g2655276"
/translation="MRPAAAGEAGAGLRWLGLAALLAALIGTPCAAHHEDKAISVPDH

```

GFCQPI S I P L C T D I A Y N Q T I L P N L L G H T N Q E D A G L E V H Q F Y P L V K V Q C S A E L K F F L C S
M Y A P V C T V L E Q A I P C R S L C E R A Q G C E A L M N K F G Q W P E R L R C E N F P V H G A G E I C V G
Q T D A P P G P G G A G R L M A Q T A G Y L P D L T P P Q A A G S F S C P R Q L K V P Y L G Y R F
L G E R D C G A P C G P G R N G L M Y F K E A E V R F A R L W G V N S V L C C A T F R T V L T Y L V D M R R F
S Y P E R P I I F L S G C Y F M A V A Y A A G F L L E E R V V C L E R S E D G Y T V A Q G T K K E G T I L F
M I L Y F E G M A S S I W V I L S L T W L A A G M K W G H E A I E A N S Q Y F H L A A W A P A V K T I T I L A
M G Q V D G D V L S G V C Y G I Y S V D S L R G F V L A P L F V Y L F I C T S F L L A G F V S L F R I T I M K H
C P T K T E K L E K L M V R I G V F S V L Y T V P A T I V A C Y F Y E Q A F R S T W E K T W L L Q T C K T Y A V P
D G S H F A P M S P D F T V E M I K Y L M T M I V G I T T G F W I W S G K T L Q S W R R F Y H R L S T G S K G E T A
Vⁿ

BASE COUNT	278 a	585 c	521 g	343 t
------------	-------	-------	-------	-------

ORIGIN

Query Match 2.1%; Score 42; DB 16; Length 1727;
Best Local Similarity 61.4%; Pred. No. 2.95e-07;
Matches 154; Conservative 0; Mismatches 94; Indels 3; Gaps 3;

Db	207	CTGCCCCAACCTGCTGGGCCACACCAACCAGGAGGACCGCAGGCCCTGGAGGTGCACCAAGTTC	266
QY	433	CTGCCCCAACCTGCTGGGCCACGAGACCATGAAGAGGTGCTGGAGCAGGCCGCGCTTGG	492
Db	267	TACCCGCTGGTCAAGTGCAGTGCTCGGCCGAGCTCAAGTTCTTCCCTCTGCTCCATGTAC	326
QY	493	ATCCCGCTGGTCATGAAGCAGTGCCACCCGGACACCAAGAAGTTCCCTGTGCTCGCTCTTC	552
Db	327	GCACCGGTTTGCACCG-T-AC-TGGAGCAGGCCCATCCACCCCTGCCGCTCCCTATGTGAG	383
QY	553	GCCCCGCTCTGCCCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCGCTCTGCGTG	612
Db	384	CGGGCTCGCCAGGGTTGTGAGCGCTCATGAACAAGTTTGGCTTCCAGTGGCCGGAGCGG	443
QY	613	CAGGTGAAGGACCGCTGCGCCCCGGTCAATGTCGCCCTTCGGCTTCCCTTGCCCGACATG	672
Db	444	CTCCGCTGCGA	454
QY	673	CTTGAGTGCGA	683

RESULT	28
LOCUS	HUMFRIZ 1923 bp mRNA PRI 24-MAY-1996
DEFINITION	Human frizzled gene product mRNA, complete cds.
ACCESSION	L37882
VERSION	g736678
KEYWORDS	frizzled gene; polarity gene.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1923) Zhao,Z., Lee,C.C., Baldini,A. and Caskey,C.T. A human homologue of the Drosophila polarity gene frizzled has been identified and mapped to 17q21.1 Genomics 27 (2), 370-373 (1995)
JOURNAL	Genomics 27 (2), 370-373 (1995)
MEDLINE	96044450
FEATURES	Location/Qualifiers 1..1923 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="fz-2" /tissue_type="ovary" /chromosome="17" /map="17q21.1" 133..1830 /gene="frizzled" 133..1830 /gene="frizzled" /function="development" /note="polarity gene; putative" /codon_start=1 /db_xref="PID:g736679" /translation="MRPRSAIPRLPLLLPAAAGPAQFHGKGISIPDHGFCOPISI PLCTDIAYNQTIMPNLGTNQEDAGLEVHQFYPLVKVQCSPELRFFLCSMYAPVCTV LEOAIPPCRSICERAROGCEALMNKFGEOWPERLRCEHFPRHGAFQICVGONHSFEDGA
gene	
CDS	

LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete
sequence.
ACCESSION AC005369
NID 93367505
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 74371)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Sequencing of human chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 74371)
AUTHORS Ricke,D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 74371)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
source Location/Qualifiers
1. 74371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5q"
/clone="119j3"
/chromosome="5"
/note="LBNL H175"
893. 1030
/rpt_family="Alu"
2295. 2438
/rpt_family="Alu"
2818. 2859
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
misc_feature join(3246. 3410,3721. 3828)
/standard_name="RLF"
/note="65% & 69% protein identity GenPept:U22377"
3431. 3724
/rpt_family="Alu"
3707. 3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
4366. 4661
/rpt_family="Alu"
5327. 5602
/rpt_family="Alu"
6586. 6956
/rpt_family="L1"
6647. 6684
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
7113. 7373
/rpt_family="Alu"
complement(7830. 8185)
/standard_name="possible repeat"
8258. 8503
/rpt_family="Alu"
9070. 9387

repeat_region
complement(9740. 9845)
/rpt_family="MER42"
complement(10440. 11015)
/rpt_family="Alu"
11950. 12250
/rpt_family="Alu"
12057. 12085
/note="(A)29"
/rpt_type=tandem
/rpt_unit=A
12365. 12645
/rpt_family="Alu"
13727. 13750
/note="(AC)12"
/rpt_type=tandem
/rpt_unit=AC
13783. 14024
/rpt_family="L1"
14175. 14470
/rpt_family="Alu"
complement(14906. 15259)
/standard_name="possible repeat"
15300. 15613
/rpt_family="Alu"
16671. 16690
/note="(T)20"
/rpt_type=tandem
/rpt_unit=T
complement(16675. 16977)
/rpt_family="Alu"
complement(16993. 17085)
/rpt_family="MER42"
complement(17678. 18276)
/rpt_family="Alu"
19305. 19583
/rpt_family="Alu"
19914. 19945
/note="(GTTT)8"
/rpt_type=tandem
/rpt_unit=GTTT
complement(19943. 20222)
/rpt_family="Alu"
21202. 21496
/rpt_family="Alu"
21736. 22035
/rpt_family="Alu"
22017. 22038
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
22321. 22457
/rpt_family="Alu"
complement(22638. 22981)
/standard_name="possible repeat"
23022. 23326
/rpt_family="Alu"
23473. 23761
/rpt_family="Alu"
23744. 23767
/note="(A)24"
/rpt_type=tandem
/rpt_unit=A
complement(24639. 24694)
/rpt_family="MER42"
complement(25349. 25713)
/rpt_family="Alu"
complement(25727. 26471)
/rpt_family="Alu"
27191. 27477
/rpt_family="Alu"
27774. 28057
/rpt_family="Alu"

QY 746 CCACCGAGGAGCTCCAAAGGTATGTGAAGCCTGCAGAAATAAAATGATGATGACAAC 804

RESULT 33
LOCUS I28278 215 bp DNA PAT 30-OCT-1996
DEFINITION Sequence 5 from patent US 5569830.
ACCESSION I28278
NID g1819054

KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 215)
AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease

JOURNAL Patent: US 5569830-A 5 29-OCT-1996;
FEATURES Location/Qualifiers
source 1. .215
/organism="unknown"

BASE COUNT 15 a 8 c 25 g 26 t 141 others
ORIGIN

Query Match 1.5%; Score 31; DB 17; Length 215;
Best Local Similarity 11.0%; Pred. No. 1.35e-01;
Matches 22; Conservative 82; Mismatches 95; Indels 1; Gaps 1;

Db 4 VTMSSSVVSRASNDKAKKDGNTTSSWTTDCNRTWGVCDTDTTYRVNDSGHNKYSS 63
QY 382 ATCCCGGCCAACCTGCAGCTGTGCCAGCGCATCGAATACCAGAACATGCGGCTGCCCAAC 441

Db 64 ANYNYGNNVGAARKHYTHTVNSGADSKTVTDSYNASGTSSTSSNGGTDGNRSGADSYGSS 123
QY 442 CTGCTGGGCCACGAGACCATGAAGGAGGTCTGGAGCAGCGCGGCTTGGATCCCGCTG 501

Db 124 KTAMTSRNRRTGKTANNVDSRNMGDASVSGDKNTKKHAKNSADGKVGSKNNGDRNNRYGT 183
QY 502 GTCATGAAGCAGTGCACCCCGACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCC-CGT 560

Db 184 GTKSNVSNCCGGNKRDPVSS 203
QY 561 CTGCCTCGATGACCTAGACG 580

RESULT 34
LOCUS OFU89259 354 bp DNA INV 14-MAR-1997
DEFINITION Oxytricha fallax 57kD zinc finger/protein chimera gene, partial cds.

ACCESSION U89259
NID g1881675
KEYWORDS .
SOURCE Oxytricha fallax.

ORGANISM Oxytricha fallax
Eukaryotae; mitochondrial eukaryotes; Alveolata; Ciliophora; hypotrichs; Stichotrichida; Oxytricha.

REFERENCE 1 (bases 1 to 354)
AUTHORS Doak,T.G., Doerder,F.P., Jahn,C.L. and Herrick,G.
TITLE A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common 'D35E' motif

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (3), 942-946 (1994)
MEDLINE 94134747
REFERENCE 2 (bases 1 to 354)
AUTHORS Witherspoon,D.J., Doak,T.G., Williams,K., Seger,J. and Herrick,G.
TITLE Selection on the protein-coding genes of the TBE1 family of transposable elements in the ciliates Oxytricha fallax and O. trifallax

FEATURES
source Location/Qualifiers
1. .354
/organism="Oxytricha fallax"
/strain="9D1"
/db_xref="taxon:5944"
/transposon="TBE1"
/note="this is a bulk sequence that was generated from a PCR product that represents many transposon templates"
<1. .>354
/codon_start=1
/transl_table=6
/product="57kD zinc finger/protein chimera"
/db_xref="PID:g1881676"
/translation="HTRDLXKHLKKAHKKXXEXEXXXXXLXLXKPKAREXXXXXXXXXX
QAXEXXVXXRXNXLXSEXTKIMIKIYKKIPVLAQIDLTSLQSYLLEDSFDKKVI
XDHQYAKVYFPNYL"
BASE COUNT 106 a 42 c 41 g 54 t 111 others
ORIGIN

Query Match 1.5%; Score 30; DB 14; Length 354;
Best Local Similarity 25.0%; Pred. No. 4.02e-01;
Matches 22; Conservative 37; Mismatches 29; Indels 0; Gaps 0;

Db 66 WGAHRHSARYTRAAGYTDCTCAAHAAAGAAAGCDAGRGARYHHYHDWARGWBGYDGW 125
Cp 1511 TGAAATCAGCTGACTCAAAACAAACAAACACCAACCAACCAAGTCACAGTTGCA 1452

Db 126 GCYYKAAYAAGCWAGMGARWYASWRTRY 153
Cp 1451 CCTATCAAACTAGCTTTAAAGTGAGC 1424

RESULT 35
LOCUS ATHSC701 4297 bp DNA PLN 26-JAN-1994
DEFINITION A.thaliana hsc70-1 gene.

ACCESSION X77199
NID g450879
KEYWORDS heat shock cognate 70-1; hsc70 gene.
SOURCE thale cress.

ORGANISM Arabidopsis thaliana
Eukaryotae; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 4297)
AUTHORS King,K.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4297)
AUTHORS King,K.
TITLE Direct Submission

JOURNAL Submitted (22-JAN-1994) K. King, Dept of Genetics, Universitaet Tuebingen, Auf der Morgenstelle 28, 72076 Tuebingen, FRG
FEATURES Location/Qualifiers
source 1. .4297
/organism="Arabidopsis thaliana"
/strain="Ostwestfalen"
/db_xref="taxon:3702"
/tissue_type="leaf"
/clone="Athsc70g"
1. .2238
/gene="Hsc70-1"
<1. .171
/gene="Hsc70-1"
/number=1
join(<1. .171,496. .2238)
/gene="Hsc70-1"
/codon_start=1
/product="heat shock cognate 70-1"
/db_xref="PID:g450880"
/db_xref="SWISS-PROT:P22953"
/translation="GTTYSCVGVWQHDRVETIIANDQGNRTTPSYVAFDSEKRLGLA
KNOVAMSPVNTVFDKRLIGRFSDDSSVQSDMKLWPFKIQAGPADKPMIYVEYKSEK
EFAAEIISSMVLIKMREIAEAYLGVTIKNAVVTVPAYFNDSQRQATKDAGVIAGLNVI
RIINEPTAAAIAYGLDKKATRVGEKNVLIFDLGGGTFDVSLLTIEGIFEVKATAGCT

gene
exon
CDS


```

/organism="Homo sapiens"
137. .249
primer_bind 137. .161
primer_bind complement(230..249)
BASE COUNT 106 a 84 c 70 g 140 t
ORIGIN
Query Match 1.4%; Score 28; DB 24; Length 400;
Best Local Similarity 80.4%; Pred. No. 3.33e+00;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 255 ATATATATTTTAAAGAGGCTAGAGTGCTTTTAGCCTTTTAAAA 300
||||| ||||| ||| | || |||| |||| ||||| |||||
QY 1794 ATATAAATTTATAATTGACTTGAGTACTTTAAGCCTTGTTAAAA 1839

RESULT 41
LOCUS ATCOR47 1073 bp RNA PLN 02-MAR-1992
DEFINITION A.thaliana cor47 mRNA.
ACCESSION X59814
NID 916228
KEYWORDS ABA responsive gene; cold-regulated gene; cor47 gene.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1073)
AUTHORS Thomashow,M.F.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-1991) M.F. Thomashow, Dept. Crop & Soil Science,
Dept. of Microbiology, Michigan State University, East Lansing, MI
48824, USA
REFERENCE 2 (bases 1 to 1073)
AUTHORS Gilmour,S.J., Artus,N.N. and Thomashow,M.F.
TITLE cDNA sequence analysis and expression of two cold-regulated genes
of Arabidopsis thaliana
JOURNAL Plant Mol. Biol. 18 (1), 13-21 (1992)
MEDLINE 92119220
COMMENT cor47 product has homology with group II LEA (late embryogenesis
abundant) proteins. Lea proteins are responsive to ABA and water
stress.
FEATURES
source Location/Qualifiers
1..1073
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="lambda ZAP"
/clone="pPH7.2"
1..887
/gene="cor47"
<1..887
/gene="cor47"
/note="Cold and ABA regulated gene"
/codon_start=3
/db_xref="PID:g388259"
/db_xref="SWISS-PROT:P31168"
/translation="NHLKATQVLKFSHFIFYLLYKLLKSRLTMAEYKNNVPEHETP
TVATEESPATTTVEVDRLGDFLGKKEEVKPOETTLESEFDHKAQISEPELAAEHE
EVKENKITLLEELQEETDEENKPSVIEKLHRSNSSSSSSDEEGEKKKKKIVE
GEEDKKGLVEKIKEKLPGHDKTAEDDVPVSTPLVTEHPVEPTTLPVEHPPEKKGILEKIKLPG
IKEKLPGHDEKAEDSPAVTSTPLVTEHPVEPTTLPVEHPPEKKGILEKIKLPG
YHAKTTEEVKKEKESDD"
BASE COUNT 365 a 188 c 276 g 244 t
ORIGIN
Query Match 1.4%; Score 29; DB 19; Length 1073;
Best Local Similarity 79.6%; Pred. No. 1.17e+00;
Matches 39; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 1025 TCTTTTATATAAGTTTGCATATGTTTATACCTTANAAAAA 1073
||||| ||||| ||| ||| ||| ||| ||||| |||||
QY 1970 TCTTGTATATAAACTTTACATCCTGTTTACCTAAAAA 2018
```

```

RESULT 42
LOCUS AF045766 1461 bp mRNA ROD 02-JUL-1998
DEFINITION Mus musculus orphan G protein-coupled receptor (GPR33) mRNA,
complete cds.
ACCESSION AF045766
NID g3282841
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1461)
AUTHORS Marchese,A., Nguyen,T., Malik,P., Xu,S., Cheng,R., Xie,Z.,
Heng,H.H., George,S.R., Kolakowski,L.F. Jr. and O'Dowd,B.F.
TITLE Cloning genes encoding receptors related to chemoattractant
receptors
JOURNAL Genomics 50 (2), 281-286 (1998)
MEDLINE 98317543
REFERENCE 2 (bases 1 to 1461)
AUTHORS O'Dowd,B.F.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1998) Department of Pharmacology, University of
Toronto, 8 Taddle Creek Rd., Toronto, ON M5S 1A8, Canada
FEATURES
source Location/Qualifiers
1..1461
/organism="Mus musculus"
/db_xref="taxon:10090"
1..1461
/gene="GPR33"
160..1179
/gene="GPR33"
/note="related to chemoattractant receptors"
/codon_start=1
/product="Orphan G protein-coupled receptor"
/db_xref="PID:g3282842"
/translation="MDLINSSTHVINVSTSLTNSTGVPTPAPKTIIAASLEMAFIIGV
ISNGLYLWMLQFKMQRTVNTLLFFHLILSYFISTLILPMATSFLODNHWVFGSVLCK
AFNSTLSVMFASVFEFLSAISVARYYLILHPVWSQQHRTPHWASRIALQIWISATILS
IPYLVERTTTHDDHKGRICKQNNYIVSTDWESKEHQTLGQWIHAACFVGRFLLGFLLPF
LVIIFCYKRVATMKKEKGLFKSSKPKVMVTAVISFFVCWMPYHVHSGLVLTKSQPLP
LHLTLGLAVVTIISFNTVVPVLYLFTGENFKVKKSILALFNSTFSDISSTERTQTLN
SETEI"
BASE COUNT 390 a 365 c 274 g 432 t
ORIGIN
Query Match 1.4%; Score 28; DB 22; Length 1461;
Best Local Similarity 73.3%; Pred. No. 3.33e+00;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db 1367 ATAAATTTATATTTTAAAAATAAACCTCTGAATCCTTAAATAAAAAA 1426
||||| ||||| || | ||||| | |||| | ||||| |||||
QY 1957 ATAAACATTATATCTTGTATATAAACCTTACATCCTGTTTACCTAAAAA 2016

RESULT 43
LOCUS A07367 1466 bp DNA PAT 18-AUG-1993
DEFINITION M.musculus VAC-alpha cDNA.
ACCESSION A07367
NID g412270
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 1466)
AUTHORS .
JOURNAL Patent: DE 3810331-A 3 05-OCT-1989;
FEATURES Location/Qualifiers
source 1..1466
/organism="Mus musculus"
```


REFERENCE 1 (bases 1 to 1567)
AUTHORS Saino,Y., Iwasaki,A. and Suda,M.
TITLE Nucleic acids encoding human placental coagulation inhibitor (PCI)
JOURNAL Patent: US 5591633-A 2 07-JAN-1997;
FEATURES Location/Qualifiers
source 1..1567
/organism="unknown"
BASE COUNT 430 a 332 c 358 g 447 t
ORIGIN
Query Match 1.4%; Score 28; DB 17; Length 1567;
Best Local Similarity 80.4%; Pred. No. 3.33e+00;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 1353 ATATATATTTTAAAGAGGCTAGACTGCTTTTAGCCTTTTAAAAA 1398
||||| ||||| ||| | || |||| |||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGACTTGAAGCCTTGTTAAAAA 1839
RESULT 47
LOCUS A13845 1575 bp DNA PAT 25-JAN-1994
DEFINITION H.sapiens PP4 mRNA.
ACCESSION A13845
NID 9490121
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1575)
AUTHORS Grundmann,U., Abel,K.J. and Kuepper,H.
TITLE Production of the anticoagulant protein PP4 by genetic engineering
JOURNAL Patent: EP 0318703-A 6 07-JUN-1989;
BEHRINGWERKE Aktiengesellschaft
FEATURES Location/Qualifiers
source 1..1575
/organism="Homo sapiens"
/db_xref="taxon:9606"
107..1069
/codon_start=1
/product="PP4"
/db_xref="PID:g490122"
/translation="MAQVLRGTVTDFPGFDERADAQTLRKAMKGLGTDDEESILTLT
RSNAQRQEISAAFKTLFGRDLLDLKSELTKFEKLI VALMKPSRLYDAYELKHALKG
AGTNEKVLTEIIASRTPEELRAIKQVYEEYEGSSLEDDVVGDTSGYQRMVLVLLQAN
RDPDAGIDEAQVEQDAQALFQAGELKWGTDEKFITIFGTRSVSHLRKVFDKYMTISG
FQIETIDRETSGNLEQLLAVVKSIRSIPAYLAETLYYAMKGAGTDDHTLIRVMVSR
SEIDLFNIRKEFRKNFATSLYSMIKGTSGDYKKALLLCCGEDD"
BASE COUNT 461 a 322 c 351 g 441 t
ORIGIN
Query Match 1.4%; Score 28; DB 17; Length 1575;
Best Local Similarity 80.4%; Pred. No. 3.33e+00;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 1324 ATATATATTTTAAAGAGGCTAGAGTGCTTTTAGCCTTTTAAAAA 1369
||||| ||||| ||| | || |||| |||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTAAAAA 1839
RESULT 48
LOCUS HUMPA4 1582 bp mRNA PRI 11-AUG-1995
DEFINITION Human placenta anticoagulant protein PP4 mRNA, complete cds.
ACCESSION M19384 J03264
NID g189614
KEYWORDS anticoagulant protein; thromboplastin inhibitor.
SOURCE Homo sapiens (clone: PP4-[14,20,26,48].) placenta cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1582)
AUTHORS Grundmann,U., Abel,K.J., Bohn,H., Lobermann,H., Lottspeich,F. and
Kupper,H.

TITLE Characterization of cDNA encoding human placental anticoagulant
protein (PP4): homology with the lipocortin family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (11), 3708-3712 (1988)
MEDLINE 88234495
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by U.Grundman, 19-APR-1988.
FEATURES Location/Qualifiers
source 1..1582
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PP4-[14,20,26,48]."
/tissue_type="placenta"
<1..1582
/note="PP4 mRNA"
151..1113
/codon_start=1
/product="anticoagulant protein 4"
/db_xref="PID:g189615"
/translation="MAQVLRGTVTDFPGFDERADAETLRKAMKGLGTDDEESILTLT
RSNAQRQEISAAFKTLFGRDLLDLKSELTKFEKLI VALMKPSRLYDAYELKHALKG
AGTNEKVLTEIIASRTPEELRAIKQVYEEYEGSSLEDDVVGDTSGYQRMVLVLLQAN
RDPDAGIDEAQVEQDAQALFQAGELKWGTDEKFITIFGTRSVSHLRKVFDKYMTISG
FQIETIDRETSGNLEQLLAVVKSIRSIPAYLAETLYYAMKGAGTDDHTLIRVMVSR
SEIDLFNIRKEFRKNFATSLYSMIKGTSGDYKKALLLCCGEDD"
variation 214
/note="g in PP4-[14,26,48]; c in PP4-20"
/replace="c"
BASE COUNT 433 a 336 c 364 g 449 t
ORIGIN 82 bp upstream of SmaI site.
Query Match 1.4%; Score 28; DB 21; Length 1582;
Best Local Similarity 80.4%; Pred. No. 3.33e+00;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 1368 ATATATATTTTAAAGAGGCTAGACTGCTTTTAGCCTTTTAAAAA 1413
||||| ||||| ||| | || |||| |||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGACTTGAAGCCTTGTTAAAAA 1839
RESULT 49
LOCUS HUMENN 1592 bp mRNA PRI 07-NOV-1994
DEFINITION Human endonexin II mRNA, complete cds.
ACCESSION J03745
NID g182111
KEYWORDS Ca2+ -dependent phospholipid binding protein; endonexin.
SOURCE Human placenta, cDNA to mRNA, (library of Clonetech Laboratories
Inc.).
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1592)
AUTHORS Kaplan,R., Jaye,M., Burgess,W.H., Schlaepfer,D.D. and Haigler,H.T.
TITLE Cloning and expression of cDNA for human endonexin II, a Ca2+ and
phospholipid binding protein
JOURNAL J. Biol. Chem. 263 (17), 8037-8043 (1988)
MEDLINE 88228020
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by H.T.Haigler, 06-APR-1988.
FEATURES Location/Qualifiers
source 1..1592
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="4q28-q32"
<1..1592
/note="endonexin II mRNA"
160..1122
/gene="ANX5"
160..1122
/gene="ANX5"
/note="ANX5"
/codon_start=1
/db_xref="GDB:G00-120-555"
/db_xref="PID:g182112"

AUTHORS

Blattner,F.R., Plunkett III,G., Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B. and Shao,Y.
The complete genome sequence of Escherichia coli K-12
Science 277 (5331), 1453-1474 (1997)
97426617

REFERENCE

2

(bases 1 to 10934)
Blattner,F.R.
Direct Submission
Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE

3

(bases 1 to 10934)
Blattner,F.R.
Direct Submission
Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

COMMENT

The E. coli K-12 sequence and its annotations have been updated. All of the ambiguous residues in our original submission have been resolved, and mis-assemblies in two repetitive regions have been realigned. The annotations have been improved and updated as well. With this release we begin designating a version number for the annotated sequence, to assist in keeping track of corrections, updates, and other changes. This is version M52 (SEPT. 02, 1997). In addition, a revised notation has been instituted which assigns each gene (protein- or RNA-encoding) a unique numeric identifier beginning with a lowercase 'b' (in the '/label' field); this will remain constant through further updates, gene identifications, etc. This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K-12 strain MGL655.
Predicted open reading frames were determined using GeneMark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332.
e-mail: mark@amber.gatech.edu

Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>).

FEATURES

source

Location/Qualifiers
1. .10934
/organism="Escherichia coli"
/strain="K-12"
/sub_strain="MGL655"
/db_xref="taxon:562"
393. .1106
/gene="yjbP"
393. .1106
/gene="yjbP"
/note="o237; sequence change joins two ORFs relative to earlier version; 99.2 pct identical to to the conceptual ORF YJBP_ECOLI SW: P32697"
/codon_start=1
/transl_table=11
/label=b4055
/product="hypothetical 26.1 kD protein in tyrB-uvrA intergenic region"
/db_xref="PID:g2367341"
/translation="MRKITQAIASVCLLFALNSSAVALASSPSPLNPGTNVARLAEQA PIHWVSQAQIENSLAGRPMPAVGFDIDDTVLFSSPGFWRGKKTFSPESEDYLNKPNVFW EKMNNGWDEFSIPKEVARQLIDMHVRGDAIFFVTGRSPTKTETVSKTLADNFHIPAT

gene

CDS

gene

CDS

NMNPVIFAGDKPQNTKSQWLQDKNIRIFYGSDNDITAARDVGARGIRILRASNSTY
KPLPQAGAFGEVIVNSEY"
1217. .1633
/gene="yjbQ"
1217. .1633
/gene="yjbQ"
/note="o138; 100 pct identical amino acid sequence and equal length to YJBQ_ECOLI SW: P32698"
/codon_start=1
/transl_table=11
/label=b4056
/product="hypothetical 15.7 kD protein in tyrB-uvrA intergenic region"
/db_xref="PID:g1790491"
/translation="MWYQKTLTLSAKSRGFHLVTDIILNQLADMPRVNIGLLHLLLOH
TSASLTLENCDPTVRHDMERFFLRTVPDNGVYEHDIYEGADMPSHIKSSMLGTSLVL
PVHKGRIQTGTWQGIWLGEHRIHGSSRRIATIQQE"
1637. .1993
/gene="yjbr"
1637. .1993
/gene="yjbr"
/note="o118; 100 pct identical amino acid sequence and equal length to YJBR_ECOLI SW: P32699"
/codon_start=1
/transl_table=11
/label=b4057
/product="hypothetical 13.4 kD protein in tyrB-uvrA intergenic region"
/db_xref="PID:g2367342"
/translation="MTISELLOYCMAPGAEQSVHNDWKATQIKVEDVLFAMVKEVEN
RPAVSLKTSPELAELLRQCHSDVRPSRHLNKAHWSTVYLDGSLPDSQIYYLVDA5YQQ
AVNLLPEKRXLLVQL"
complement(2028. .4850)
/gene="uvrA"
complement(2028. .4850)
/gene="uvrA"
/note="o940; 99 pct identical amino acid sequence and equal length to UVRA_ECOLI SW: P07671; CG Site No. 21"
/codon_start=1
/transl_table=11
/label=b4058
/product="excision nuclease"
/db_xref="PID:g2367343"
/translation="MDKIEVRGARTHNLKNINLVIPRDKLIVVTGLSGSGKSSLAFTD
LYAEQRRYVESLSAYARQFLSLMEKPDVDHIEGLSPAISIEQKSTSHNPRSTVGTIT
EIHLYRLLEFARVGEPRCPDHDVPLAAQTVSQMVDNVLSQPEGKRLMLLAPIKERKG
EHTKTLENLASQGYIRARIDGVECDLSDPPKLELQKKHTIEVVVDRFKVRDLDLQRLA
ESFETALELSGTAIVADMDDPKAEELLESANFACPICGYSMRELEPRLFSFNNPAGA
CPTCDGLVQOYFDPDRVIONPELSLAGAIRGWDNRNFYFQMLKSLADHYKFDVEA
PWGSLSANHKVVLVYSGKENIEFKYMNDRGDTSIRRHPPFEGVLHNMERRYKETESSA
VREELAKFISNRPCASCEGTRLRREARHVVENTPLPAISDMSIGHAMEFFNNLKLKAG
QRAKIAEKILKEIGDRLLKFLNVNGLNYLTLRSAAETLSGGEAQIRLASIGAGLVGV
MYVLDEPSIGLHQRDNERLLGTLIHLRDLGNTVIVVEHDEDAIRAADHVIDIGPAGV
HGGEVVAEGPLEAIMAVPESLTQYMSGKRKIEVPKKRVSPANPEKVLKLTGARGNNLK
DVTLTLPVGLFTCITGVSGSKSTLINDTLEPIAQRLNGATIAEPAPYRDIQGLEHF
DKVIDIDQSPIGRTPRSNPATYTGFTPVRELFAGVPESRARGYTPGRFSFNVRGRC
EACQGDGVIKVEMHFLPDIYVPCDQCKGRYNRETLEIKYKGTIHEVLDMTIEEARE
FFDAVPALARKLQTLMDVGLTYIRLQGSATLSGGEAQVKLARELSKRGTTGQTLIYL
DEPTTGLHFADIQQLLDVLHKLKRDQGNTIVVIEHNLDVIKTADWIVDLGPPEGSGGGE
ILVSGTPETVAECEASHTARFLKPL"
5104. .5640
/gene="ssb"
5104. .5640
/gene="ssb"
/note="o178; CG Site No. 150"
/codon_start=1
/transl_table=11
/label=b4059
/product="single-strand DNA-binding protein"
/db_xref="PID:g1790494"
/translation="MASRGVNKVLVGNLQDPEVRYMPNGGAVANITLATSESWRDK
ATGEMKEQTEWHRVVLFGKLAEVASEYLRKGSQVYIEQLRTRKKTQSGQDRYTFEV
VNVVGGTMQLGGRQGGGAPAGGNIGGGQPGQGWGQPPQPPQGGNQFSGGAQSRPQOSA

gene

CDS

gene

CDS

gene

CDS

gene PAAPSNPEPPMEDDDIPF"
complement(5739. .6089)
/gene="yjcb"
CDS complement(5739. .6089)
/gene="yjcb"
/note="f116; 100 pct identical amino acid sequence and
equal length to YJCB_ECOLI SW: P32700"
/codon_start=1
/transl_table=11
/label=b4060
/product="hypothetical 13.0 kD protein in ssb-soxs
intergenic region"
/db_xref="PID:gl790495"
/translation="MYHASREQRRAYQGIGSFQRCNCLMATLTGTGVLLRWQLLSAVMM
FLASTLNIRFRSDYVGLAVISSGLGVVSACWFAMGLLIGITMADITAIWHNIESVMIE
EMNQTTPQWPMILT"
6450. .8036
/gene="yjcc"
CDS 6450. .8036
/gene="yjcc"
/note="o528; 100 pct identical to YJCC_ECOLI SW: P32701;
similar to Azorhizobium caulinodans hypoth. protein, ntrC
3' region"
/codon_start=1
/transl_table=11
/label=b4061
/product="hypothetical 60.8 kD protein in ssb-soxs
intergenic region"
/db_xref="PID:gl790496"
/translation="MSHRARHQLLALPGIIFLVLPFIILSLWIAFLWAKSEVNNQLRT
FAQLALDKSELVIRQADLVSDAAERYQGVCTPAHQKRLMNIIRGYLINELIYARDN
HFLCSSLIAPVNGYTIAPADYKREPNSIYYRDTPFFSGYKMTYMQRGNYVAVINPL
FWSEVMSDDPTLQWGYDVTVTKTFSLSKEASATFSPLIHLKDLTVQRNGLYATVY
STKRPIAAIVATSQRLITHEYNHLIFALPAGLSLVLLLLLWLRIRQNYLSPRKLQ
RALEKHQLCLXYQIPLIDIKTEKCIGAEALLRWPGQGIIMNPAEFIPLAEKEGMIEQI
TDYVIDNVFRDLGDLATHADRYVSINLSASDFHSRLIARINQKTEQYAVRPQQIKF
EVTEHAFLDVDMPTPIILAFRQAGYEVAIDDFGIGYSNLHLNLSLNVDLIKDKSFVE
TLTHKTSHLIAEHIELAHSLSGLKLTIAEGVETEBOVNWLRKRGVRYCQGWFFAKAMP
PQVFMQWMEQLPARELTRGQ"
complement(8039. .8362)
/gene="soxs"
CDS complement(8039. .8362)
/gene="soxs"

....
Note: remainder of annotations omitted.

Query Match 1.4%; Score 28; DB 12; Length 10934;
Best Local Similarity 71.9%; Pred. No. 3.33e+00;
Matches 46; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Db 3127 CCTTCTCGGACCACCTTCACCGCCGTGAACACCTCGCCCGGCCCAATGTCGATCATGG 3186
||||||| ||| ||||||| |||| ||||| |||| |||| |||| |||| ||||
Cp 1090 CCTTGTCCACCGCTTCACCGAGGTGATCACCAGCTCCCCACCCCTGTTCTGTCCCATGA 1031

Db 3187 TCAG 3190
|||
Cp 1030 CCAG 1027

RESULT 58
LOCUS AFF28J12 110102 bp DNA PLN 03-FEB-1998
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone F28J12 (ESSAII
project).
ACCESSION AL021710
NID g2832639
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
eudicotyledons; Rosidae; Capparales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 110102)
AUTHORS Bevan,M., Hilbert,H., Braun,M., Holzer,E., Brandt,A.,

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source

exon

CDS

gene

CDS

exon

gene

intron

exon

intron

exon

Duesterhoeft,A., Bancroft,I., Mewes,H.W., Mayer,K. and Schueller,C.
Unpublished
2 (bases 1 to 110102)
EU Arabidopsis sequencing,project.
Direct Submission
Submitted (03-FEB-1998) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bsrc.ac.uk

Location/Qualifiers
1. .110102
/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
1988. .3739
/gene="F28J12.10"
/number=1
1988. .3739
/gene="F28J12.10"
/note="strong similarity to neoxanthin cleavage enzyme,
Lycopersicon esculentum, PATCHX:E325797; and
viviparous-14, Zea mays, PATCHX:G2232017"
/codon_start=1
/product="neoxanthin cleavage enzyme - like protein"
/db_xref="PID:el249600"
/db_xref="PID:g2832640"
/translation="MVSLTMPMSGGIKTWPQAQIDLGRPIKRPKVIKCTVQIDVY
ELTKKRLFTPRTTATPPQHNLRLNIFQKAAAIADAERALISHEQDPSLPKTDAP
RVQIAGNYSVPVPESSVRNLTVEGTIPDCIDGVYIRNGANPMFEPTAGHHLFDGDMV
HAVKITNGSASYACRFTKTERLVQEKRLGRPVFPAKIGELHGHSGIARLMLFYARGLC
GLINNQNGVGVANAGLVYFNNRLAMSEDDLQYQKAITQIGDQTVGRYDFDQGLKSA
MIAHPKLDPVTKELHALSYDVKKPYLKYFRFSPDGKVSPELEPPTMIHDFAIT
ENFVVPDQQVVKLGEMISGKSPVFDGKVSRLGIMPKDATEASQIIWVNSPEIFC
FHLNNAWESPETEIVVIGSCMPADSIFFNERDESLRSVLSEIRNLRTKTTRRSLL
VNEDVNLIEIGMVNRNRLGRKTRFAFLAIAYPWPKVSGFAKVLDLCTGEMKKYIYGKEY
GGEPFFLPNGSNGEENDDGYIFCHVHDEETKTSELQIINAVNLKLEATIKLPSRV?
YGFHGTFFVDSNELVDQL"
1988. .3739
/gene="F28J12.10"
complement(join(5457. .5645,5970. .6074,6165. .6265,
6367. .6436,6535. .6597,6677. .6743,6822. .6946,7021. .7098,
7188. .7234,7348. .7472,7566. .7702))
/note="strong similarity to (S)-2-hydroxy-acid oxidase (Ec
1.1.3.15), peroxisomal, Spinacia oleracea (spinach),
PIR1:OXSPH; contains EST gb:T88637, N37256"
/codon_start=1
/product="glycolate oxidase - like protein"
/db_xref="PID:el249601"
/db_xref="PID:g2832641"
/translation="MEITNVMEYEKIAKELPKMVDYDYASGAEDQWTLQENRNAFSR
ILFRPRILIDVSKIDVSTTVLGENISMPIMIAPTAMQKMAHPDGEATARATSAGTI
MTLSSWATCSVEEVASTGPIRFFQLVYKDRNVVIQLVKRAEEAGFAKIALITVDTPR
LGRRESDIKNRFALPRGLTLKNEFGLDLGKIDKTNDSGLASYVAGVDQSLSWKDIKW
LQSITSLPILVKGVITAEADARIAVEYGAAGIIVSNHGARQLDYVPATIVALEEVVKAV
EGRIPVLDGGVRRGTDVFKALALGASGVFVGRPSLFLSLAADGEAGVRKMLQMLRDEF
ELTMALSGCRSLREISRTHIKTDWDTPHYLSAKL"
complement(5460. .5645)
/gene="F28J12.20"
/number=1
5460. .7702
/gene="F28J12.20"
complement(5646. .5969)
/gene="F28J12.20"
/number=1
complement(5970. .6074)
/gene="F28J12.20"
/number=2
complement(6075. .6164)
/gene="F28J12.20"
/number=2
complement(6165. .6265)

/gene="F28J12.20"
/number=3
complement(6266. .6366)
/gene="F28J12.20"
/number=3
complement(6367. .6436)
/gene="F28J12.20"
/number=4
complement(6437. .6534)
/gene="F28J12.20"
/number=4
complement(6535. .6597)
/gene="F28J12.20"
/number=5
complement(6598. .6676)
/gene="F28J12.20"
/number=5
complement(6677. .6743)
/gene="F28J12.20"
/number=6
complement(6744. .6821)
/gene="F28J12.20"
/number=6
complement(6822. .6946)
/gene="F28J12.20"
/number=7
complement(6947. .7020)
/gene="F28J12.20"
/number=7
complement(7021. .7098)
/gene="F28J12.20"
/number=8
complement(7099. .7187)
/gene="F28J12.20"
/number=8
complement(7188. .7234)
/gene="F28J12.20"
/number=9
complement(7235. .7347)
/gene="F28J12.20"
/number=9
complement(7348. .7472)
/gene="F28J12.20"
/number=10
complement(7473. .7565)
/gene="F28J12.20"
/number=10
complement(7566. .7702)
/gene="F28J12.20"
/number=11
8557. .8838
/gene="F28J12.30"
/number=1
8557. .13947
/gene="F28J12.30"
join(8557. .8838,8913. .9050,9136. .9198,9284. .9373,
9450. .9572,9668. .9755,10089. .10163,10261. .10280,
10348. .10368,12117. .12725,12806. .13321,13362. .13622,
13693. .13947)
/gene="F28J12.30"
/note="similarity to MucD - Pseudomonas aeruginosa,
PATCH:G1184684; contains EST gb:T22208"
/codon_start=1
/product="putative protein"
/db_xref="PID:el249602"
/db_xref="PID:g2832642"
/translation="MALASSKAFSSIFNTLSPINQSKFVLACSGSNHVDVIDRRRRIM
IFGSSALTSLLGSNQRLPMESAIALEQFKEKEEELEEEERNVNLFOKTSPSVVY
IEATELPKTSGLDITDEENGKIEGTSGFVWDKLGHIVTNVHIAKLATDQFGLQRC
KVSVDAGKTRFSKEGKIIVGLDPDNLAVLKIETEGRELNPVVLGTSNDLRVGQSCFA
IGNPYGYENTLTIGVSGLGREIPSPNGKSISEAIQTADADINSNGSGPLDLSYGHTI
GYNTATFTRKSGSMSSGLIETSEVLNSVYLTMTVERKKRQIQRNNSERNQKRRIIS
HDKINRDELVVYRILCPIDVVGGVIGKSGKVINAIRHNTKAKIKVFDQLHGCSQRVIT

IYCSVKEKQEEIGFTKSENEPLCCAQDALLKVYDAIVASDEENNTKTNVDRDNKECR
LLVPFSQSSSLIGKAGENIKRIRRRTRASVKVSKDVSDPSHVCAMEYDNVVVISGEP
ESVKQALFAVSAIMYKINPRENIPLDSTSQDVPAAASVIVPSDLNSVYPOTGFYSNQD
HILQQAGVPSYFNALSVDFOGYAETAAANPVPFASSLPVTHGGGSSSEELVFKV
LCPLCNIMRVIGKGGSTIKRIEASGSCIEVNDSTRKCGDDECVIIVTATEILFCCLS
TPFVFMQSPDDMKMSMAVEAVLLQEXYNDEDAENVKMQLLVSSKVICVIGKSGSVIN
EIRKRTNANICISKGKKDDLVEVSGSVSRDALIQIVLRLREDVLGDKDSVATRKPP
ARTDNCFLSGSSNAGYTLPSFMSSMASTSGFHGYGSPVKSCILDFH"
8839. .8912
/gene="F28J12.30"
/number=1
8913. .9050
/gene="F28J12.30"
/number=2
9051. .9135
/gene="F28J12.30"
/number=2
9136. .9198
/gene="F28J12.30"
/number=3
9199. .9283
/gene="F28J12.30"
/number=3
9284. .9373
/gene="F28J12.30"
/number=4
9374. .9449
/gene="F28J12.30"
/number=4
9450. .9572
/gene="F28J12.30"
/number=5
9573. .9667
/gene="F28J12.30"
/number=5
9668. .9755
/gene="F28J12.30"
/number=6
9756. .10088
/gene="F28J12.30"
/number=6

...
Note: remainder of annotations omitted.

Query Match 1.4%; Score 28; DB 19; Length 110102;
Best Local Similarity 76.9%; Pred. No. 3.33e+00;
Matches 40; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

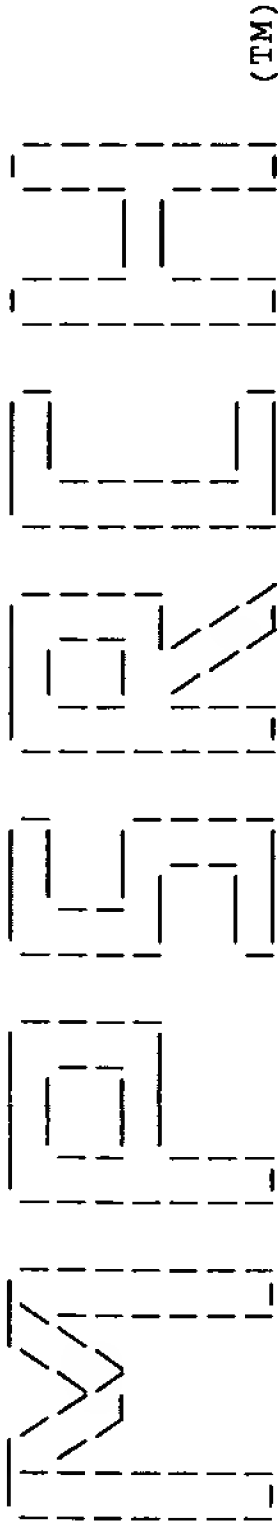
Db 72216 TTTTTTTTTTTTAGGTAAAGCATTTGTAATAATTTGATATCTATGAAT 72267
|||||
Cp 2019 TTTTTTTTTTTTAGGTAAACAGGATGTAAAGTTTATATACAAGAAT 1968

RESULT 59
LOCUS F8K4 115641 bp DNA PLN 04-AUG-1998
DEFINITION Arabidopsis thaliana chromosome 1 BAC F8K4 sequence, complete
sequence.
ACCESSION AC004392
NID g3282170
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; Capparales; Brassicaceae;
Arabidopsis.
REFERENCE 1 (bases 1 to 115641)
AUTHORS Vysotskaia,V.S., Schwartz,J.R., Toriumi,M., Kwan,A., Yu,G., Oji,O.,
Liu,S., Li,J., Araujo,R., Au,M., Brendel,V., Buehler,E.,
Conway,A.B., Conway,A.R., Dewar,K., Feng,J., Kim,C., Kurtz,D.,
Li,Y., Palm,C.J., Shinn,P., Sun,H., Davis,R.W., Ecker,J.R.,
Federspiel,N.A. and Theologis,A.
Arabidopsis thaliana chromosome 1 BAC F8K4 sequence

JOURNAL
REFERENCE 2 (bases 1 to 115641)
AUTHORS Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-1998) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE 3 (bases 1 to 115641)
AUTHORS Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1998) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE 4 (bases 1 to 115641)
AUTHORS Theologis.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1998) Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA
REFERENCE 5 (bases 1 to 115641)
AUTHORS Theologis.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-1998) Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA
COMMENT
This sequence is of BAC F8K4 from Arabidopsis thaliana chromosome 1. The sequence does not represent the sequence of the entire insert of this clone. It is shorter by 3062 bp because we submit only the unique sequence of the clone. In order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide a small overlap (200 bp) between overlapping submitted clones. The 3' end of this sequence overlaps by 200 bp the 5' end of the sequence of BAC F19K23.

FEATURES
source
1. .115641
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F8K4"
1468. .3346
/gene="F8K4.1"
join(1468. .1959,2350. .2575,2668. .2890,3055. .3346)
/gene="F8K4.1"
/note="Similar to
glucose-6-phosphate/phosphate-translocator (GPT)
gb|AF020814 from Pisum sativum."
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g3367515"
/translation="MLSSIKPSSSSFSTAISGVSRRSIPTKLKFSPLLIKNCHNQSF
NANVVSQKPLHISSASNFKEVKVEAYEADRSRPLDINIELPDEQSAQKLKIGIYFA
TWWALNVVENIYNKKVLNAFPYPWLTSLSLACGSLMMLVSWATRIADAPKTDLEFWK
TLFPVAVAHTIGHVAATVSMKSAVVSFTHIKSGEPAFSLVSRFFMGETFPPLPVLS
LLPIIGCALAAITELNFTNITGFMGAMISNLAFFVRNIFSKKGMKGSVSGMNYIACL
SMMSLVILTPFSIAVEGPMQMAAGWQNAVSQVGNFVWVVAQSVFYHLNQVSYMSL
DQISPLTFSIGNTMKRISVIVASIIIFHTPIQPVNALGAAIAIFGTFLYSQVNKQTYS
CLPTSTWPFSEQFLVLI"
6976. .9754
/gene="F8K4.2"
join(6976. .7131,7230. .7299,7426. .7484,7603. .7678,
7852. .7926,8035. .8122,8222. .8474,8546. .8661,8775. .9033,
9133. .9235,9320. .9428,9556. .9754)
/gene="F8K4.2"
/note="Similar to beta-glucosidase BGQ60 precursor
gb|L41869 from Hordeum vulgare."
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g3367516"
/translation="MKNLTSFVIVILQSLFLHFVYGRHQSSKNILVDSSPPSPDFLF
GTASSAYQEGAFLTGDKSLNNWDVFTHKNPGLDKNNADRAVDQYNRFEDIQLMS
FLGVNSYRESISWCRILPRGRFGEINVLGIKYYNIFIDALISRGIKPFTLNVHDYPQ
ELEDRFQWLNPQMOKFEGYLADICFKHFNVRKYWTTLNEPNQKQKIGILYTGKFPFS
RCSSPYGNCQGNSETEPFIAAHNMILAHAKAVNIYKTKYQKQSGSIGIVVQTSWFE
PISDSNADKEAARAQSFYSNWLDPVIYGYKPKEMVDILGPALPQFSNEVKNLEKS
RADFVGINHITSYFIQDCLTSACNTGHGAFKAEGYALKLDRKGNVTIGELTDVNWQHI
DPTGFHKMLNLYLKDRYPNMPMFITENGFGDLQKQKPTTDELLNDTKRIQYMSGYLEAL

gene
11528. .15385
/gene="F8K4.3"
join(11528. .11625,12458. .12539,12625. .12694,13043. .13101,
13216. .13291,13473. .13547,13658. .13745,13819. .14071,
14162. .14277,14372. .14630,14741. .14843,14938. .15046,
15190. .15385)
/gene="F8K4.3"
/note="Similar to F4i1.26 putative beta-glucosidase
gi|3128187 from A. thaliana BAC gb|AC004521. ESTS
gb|N97083, gb|F19868 and gb|F15482 come from this gene."
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g3367517"
/translation="MSPTISSAAATPTPSPNGSTFSAALALLVNRNRSCLHQTSDSS
PFPSDFLFGTASSAFQYEGAFLTGKGLNNWDFAHENPGKIVDGSNGDIATDOYHRY
MEDIQSMNFLGVNSYRLSISWSRVLPNGRFGVINYKGIKYNLNLIDALIKKGITPFVT
LNHFDPQELERNFKSWLSSEMOKDFGYLADICFKHFGDRVKHWITINEPNQHISLAY
RSLFFPPARCSPMPYCNCTHGNSETEPFIAAHNMILAHAKAIQIYRTKYREQKGIGI
VVQTSWFEPISDSIADKNAARAQSFYSNWLDPVYGYKYPEEMVNLGSLPKFSSN
EMNSLMSYKSDFLGINHYTSYFIQDCLITACNSGSGASKSEGLALKDRKGNVSIGEL
TDVNWQHIDPNGFRKMLNLYLKNRYHNIPMYITENGFGQLQKPETTVEELLHDTKRQY
LSGYLDALKAAAMRGANVKGYFAWSLLDNFEWLYGYKVFGLFHVDTTLKRTPKQSA
TWYKNFIEQNVNIEDQIDK"
16933. .18807
/gene="F8K4.4"
join(16933. .17047,17094. .17705,17750. .18807)
/gene="F8K4.4"
/note="Strong similarity to unknown protein T28M21.21
gi|2088661 from A. thaliana BAC gb|AF002109."
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g3367518"
/translation="MDNDRGIFHMEKDGKLSLVYRPPIIQTRPTSSGRIVIRWKNR
DFPYTIRFSSITNVFPFGSIINYRIKYAHDHDFVLPLFWCNDKEDFVNGCSCACGSY
FGTDYFCNCNMVHKECVESPLKIKHPYHPEHSELEYIHPPFAHNIECLCGRIASG
QVYCSICRVLMSHCAMKTIPMVINQPKRHDHPLIFFPKQNSLTNCVGLLRKLYST
YICAICNFVAHKDCMYSPRIIKISHCSTICROSIDSIDYGAYSCAMCGDYVVSRCAL
GKDVDGEDLENVPEEDNKLDQVTSFDIISDGVILHFLHDHHLRLKVSILYDENKFCQ
ACVLPIVEGNFYSCMEDDFIFHETCPNSCRILLHALHPHQLTLEHVEYDSCRFMCNAC
ERFCGGFIYGCHIRECKFYLDVRCAIYEPENYKHYHEHSMFLASDPKEELLCQVCKTI
CHKQLNCIKCDFIVCIKCATLPEKVRYKKDKHFLRVLFGEELCKGDWCCEVCERNRAF
DTKVFWCNECGTFHIECLFGGDLIYVKPGQTLKFFDRDAEILGKSNLSRPLCMNPD
CKNPCQVKILKTYNDIVCSVKCGNQVFY"
24330. .26774
/gene="F8K4.5"
24330. .26774
/gene="F8K4.5"
/note="Strong similarity to unknown protein T28M21.21
gi|2088661 from A. thaliana BAC gb|AF002109."
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g3367532"
/translation="MDSIGRFHETETDDKLYQQLVYVHPKYTPIPQTKSPTSSHKTAS
HDLLFQPLFLCPHPRVTQLSDSITFPISNSPEYIVSATKSDQYEDHSLPLFWCNN
KEIDVDGGCDICSGNFGTDYILCVHCDSIYHKECVQSPFKIEHPYHPEHSLQLSYLD
LNARIMECFCCGRRATYLVYCTICDVLMTVCAMKSIPLGGFHELENGKGLFFVYHH
PKYQIIPQTKSPSSDKTANHDLPFQPLFLCPHLRVTKNSSDSLTYPIISYPEYIVST
KSEDSLVLPLFWCNKEFNVDGDCICSGNSTATTDLVYFCYCDKIYHKECVQSPISI
KHPYHPEHSLQLSYRLPSAPDIECLCCGTTATDLVYGTICEAVMHTPCAMKSIPEVI
DQPKSHDHSLTLPFQASLTNCVGLLRKNKYPTFFLCRCNFVAHKDCMSPHIIKISR
HHRISYASLQSEWSCGVCRKSIDGDYGAYTCECKDYVVHVRALRKDVWDGVEL
EGVPEEDDITKDVGPEFIISEGVILHFLHDHHLQLEVSILYDEKFCQACVPIYEGN
HYICVECDFTIHEICAKDRRRIQHALHPHPLTLNAVSKYDTGDFECNACVRNSGGFVY
MCSRKECNFDLDVRCAIASEPENFQGHEHSLFLAIDSEKPIQCMCKSESHIPLNCIK
CDFVVCIKCATLPEFKAKYKHDRHSLTILWGEVCGKDWCEVCERNLGDTDTKVFWCN
ECCTTLHIECLYGEILYVLKGYIFDSGGKEFQILGKATSSRPFECNICKNRCQKGIFT
RDNRIACSLICASDIY"
33258. .39166
/gene="F8K4.6"
join(33258. .33695,33792. .33939,34028. .34126,34442. .34647,
34733. .34978,35053. .35215,35298. .35558,35630. .35751,



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Nov 5 12:14:45 1998; Maspar time 291.26 Seconds
946.855 Million cell updates/sec
Tabular output not generated.

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGCGGCCGC 2027
Comp: CTTAAGCGGAAGTACCGGA.....TTTTTTTTTTCGCCGGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 60 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.846; Variance 6.883; scale 1.431

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	201	9.9	204	20	T23535 Human gene signature	1.01e-97
2	58	2.9	2334	37	T89889 Human frizzled gene 5	6.84e-16
3	56	2.8	2421	37	T89892 Mouse frizzled gene 8	7.60e-15
4	44	2.2	91	9	Q51746 Oligonucleotide probe	9.73e-09
5	45	2.2	172	32	T76363 Human interleukin 8 a	3.11e-09
6	43	2.1	204	1	N81164 Base substituted E.co	3.03e-08
7	41	2.0	91	9	Q51746 Oligonucleotide probe	2.88e-07
8	40	2.0	100	32	T76186 Human IL4 receptor an	8.76e-07
9	40	2.0	114	12	Q70466 Generic DNA sequence	8.76e-07
10	40	2.0	114	12	Q70468 Generic DNA sequence	8.76e-07
11	40	2.0	204	1	N81164 Base substituted E.co	8.76e-07
12	39	1.9	114	12	Q70469 Generic DNA sequence	2.65e-06
13	38	1.9	114	12	Q70467 Generic DNA sequence	7.95e-06

14	1.9	114	12	Q70465	Generic DNA sequence	7.95e-06
15	1.9	114	12	Q70469	Generic DNA sequence	2.65e-06
16	1.9	114	12	Q70468	Generic DNA sequence	7.95e-06
17	1.9	114	12	Q70466	Generic DNA sequence	7.95e-06
18	1.9	178	32	T76405	Human endothelin-1 an	2.65e-06
19	1.9	190	32	T76452	Chymase antisense oli	2.65e-06
20	1.8	114	12	Q70470	Generic DNA sequence	6.96e-05
21	1.8	114	12	Q70472	Generic DNA sequence	6.96e-05
22	1.8	114	12	Q70467	Generic DNA sequence	6.96e-05
23	1.7	114	12	Q70473	Generic DNA sequence	5.87e-04
24	1.7	114	12	Q70471	Generic DNA sequence	5.87e-04
25	1.7	114	12	Q70465	Generic DNA sequence	5.87e-04
26	1.7	114	12	Q70471	Generic DNA sequence	5.87e-04
27	1.7	114	12	Q70470	Generic DNA sequence	5.87e-04
28	1.7	250	32	T76438	Substance P antisense	5.87e-04
29	1.7	317	32	T76274	Human neutrophil elas	2.03e-04
30	1.6	114	12	Q70472	Generic DNA sequence	4.74e-03
31	1.6	178	32	T76405	Human endothelin-1 an	4.74e-03
32	1.6	264	32	T76445	Substance P receptor	4.74e-03
33	1.5	88	32	T76170	Human IL3 receptor an	3.65e-02
34	1.5	168	32	T76270	Human MDNCF antisense	1.32e-02
35	1.5	190	32	T76452	Chymase antisense oli	1.32e-02
36	1.5	317	32	T76274	Human neutrophil elas	3.65e-02
37	1.5	565	6	Q35072	HCV envelope region n	3.65e-02
38	1.5	3871	2	N71302	HSV-1 gB and surround	3.65e-02
39	1.4	69	32	T76167	Human IL3 antisense o	2.67e-01
40	1.4	70	32	T76184	Human IL4 antisense o	9.94e-02
41	1.4	114	12	Q70473	Generic DNA sequence	2.67e-01
42	1.4	128	32	T76233	Human IL6 antisense o	9.94e-02
43	1.4	128	32	T76233	Human IL6 antisense o	2.67e-01
44	1.4	130	32	T76152	Human vascular cell a	2.67e-01
45	1.4	162	32	T76307	Human RANTES antisens	2.67e-01
46	1.4	172	32	T76363	Human interleukin 8 a	9.94e-02
47	1.4	200	32	T76398	Human leukotriene C4	2.67e-01
48	1.4	1454	1	N82107	PAP-I cDNA from human	2.67e-01
49	1.4	1466	2	N91353	Vascular anti-coagula	2.67e-01
50	1.4	1466	1	N80801	Sequence encoding vas	2.67e-01
51	1.4	1567	1	N81113	Plasmid contg. placen	2.67e-01
52	1.4	1575	1	N90112	Anticoagulant PP4 DNA	2.67e-01
53	1.4	1605	1	N91821	Endonexin II complete	2.67e-01
54	1.4	1840	11	Q65611	Feline zona pellucida	9.94e-02
55	1.4	2259	37	T89891	Mouse frizzled gene 7	9.94e-02
56	1.3	125	32	T76423	Human endothelin ETA	7.04e-01
57	1.3	144	32	T76416	Human endothelin rece	7.04e-01
58	1.3	145	32	T76299	Human defensin 3 anti	7.04e-01
59	1.3	789	2	Q10792	Bovine vascular endot	7.04e-01
60	1.3	961	7	Q44259	Bovine VEGF-164 codin	7.04e-01

ALIGNMENTS

RESULT 1
ID T23535 standard; cDNA to mRNA; 204 BP.
AC T23535;
DT 02-SEP-1996 (first entry)
DE Human gene signature HUMGS05382.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1393; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 204 BP; 54 A; 59 C; 35 G; 55 T;

Query Match 9.9%; Score 201; DB 20; Length 204;
Best Local Similarity 99.5%; Pred. No. 1.01e-97;
Matches 201; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 gatctcagctcccggtccccaagcacactcctagctgtcctcagttctcagcctggcagct 60
|||||
QY 1207 GATCTCAGCTCCCGTTCCTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT 1266
|||||

Db 61 tccccctgccttttgccacgtttccccaagcacactcctagctgtcctcagttctcagcctggcag 120
|||||
QY 1267 TCCCCCTGCTTTTGCACGTTTGCATCCCCAGCATTTCTCAGTTTATAGGCCACAGGAG 1326
|||||

Db 121 tggatagctgttttcacctaaggaaagccaccccgaaatcttgtagaaatattcacaact 180
|||||
QY 1327 TGGATAGCTGTTTTCACCTAAAGGAAAGCCACCCGAACTCTGTAGAAATATTCAAAC 1386
|||||

Db 181 antaaaaatcatgaatatatttta 202
|||||
QY 1387 AATAAAATCATGAATATTTTAA 1408
|||||

RESULT 2
ID T89889 standard; DNA; 2334 BP.
AC T89889;
DT 27-APR-1998 (first entry)
DE Human frizzled gene 5 (Hfz5) encoding a Wnt receptor.
KW Wnt receptor; human frizzled gene 5; Hfz5 gene;
KW signal transduction; cancer; cell growth; cell proliferation;
KW mammary tumour; oncogene; therapy; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 321..2078
/*tag= a
WO9739357-A1.
PN 23-OCT-1997.
PD 11-APR-1997; U06049.
PF 12-APR-1996; US-015307.
PR (UYJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI; 97-526631/48.
DR P-PSDB; W31271.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 36-37; 61pp; English.
CC This sequence comprises novel human frizzled gene 5 (Hfz5)
CC that encodes a transmembrane receptor, frizzled-5 (see W31771), a
CC Wnt receptor (WntR). Novel frizzled family members have been
CC identified in Drosophila, mouse, human and Caenorhabditis (see
CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
CC can be used in a novel, claimed method of screening for compounds
CC which modulate the binding of a Wnt polypeptide (secreted proteins
CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
CC involved in (mammary) cancer and other processes involving growth,

CC development and proliferation (both normal and abnormal)).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 2334 BP; 356 A; 803 C; 736 G; 439 T;

Query Match 2.9%; Score 58; DB 37; Length 2334;
Best Local Similarity 59.9%; Pred. No. 6.84e-16;
Matches 175; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Db 439 tgtgcgcgggcacgtgctacaacctgacgcacatgcccacacagttcaaacacgacacgc 498
|||||
QY 401 TGTGCCACGGCATCGAATACCGAACAATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA 460
|||||

Db 499 aggacgagggcgccctggaggtgacacagttctggccgctggtggagatccaatgctgc 558
|||||
QY 461 TGAAGGAGGTGCTGGAGCAGGGCCGGCGCTTGGATCCCCGCTGTCATGAAGCAGTGCCACC 520
|||||

Db 559 cggacctgcgctttcttcctatgcaactatgtacacgcccacatctgtctgccccgactaccaca 618
|||||
QY 521 CGGACACCAAGAAGTTCTCTGTCTCGCTCTTCGCCCCCGCTCTGCTCGATGACCTAGACG 580
|||||

Db 619 agccgctgcccgcctgcccgtcggtgtgcgagcgcgccaagcgcggtctgcgcgctga 678
|||||
QY 581 AGACCATCCAGCCATGCCACTCGCTCTGCTGTCAGGTGAAGGACCGCTGCGCCCGGTCA 640
|||||

Db 679 tgcgcagtagcggcttcgctgcccgcgagcgcgcagctgagctgcgacgcgcctccc 730
|||||
QY 641 TGTCCGCTTCGGCTTCCCTTCCCTGSCCGACATGCTTGAGTGCAGACCGTTTCCC 692
|||||

RESULT 3
ID T89892 standard; DNA; 2421 BP.
AC T89892;
DT 27-APR-1998 (first entry)
DE Mouse frizzled gene 8 (Mfz8) encoding a Wnt receptor.
KW Wnt receptor; mouse frizzled gene 8; Mfz8 gene;
KW signal transduction; cancer; cell growth; cell proliferation;
KW mammary tumour; oncogene; therapy; ds.
OS Mus musculus.
FH Key Location/Qualifiers
FT CDS 188..2245
/*tag= a
WO9739357-A1.
PN 23-OCT-1997.
PD 11-APR-1997; U06049.
PF 12-APR-1996; US-015307.
PR (UYJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI; 97-526631/48.
DR P-PSDB; W31274.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 46-47; 61pp; English.
CC This sequence comprises novel mouse frizzled gene 8 (Mfz8)
CC that encodes a transmembrane receptor, frizzled-8 (see W31274),
CC a Wnt receptor (WntR). Novel frizzled family members have been
CC identified in Drosophila, mouse, human and Caenorhabditis (see
CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
CC can be used in a novel, claimed method of screening for compounds
CC which modulate the binding of a Wnt polypeptide (secreted proteins
CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
CC involved in (mammary) cancer and other processes involving growth,
CC development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 2421 BP; 403 A; 772 C; 779 G; 467 T;

Query Match 2.8%; Score 56; DB 37; Length 2421;
Best Local Similarity 59.4%; Pred. No. 7.60e-15;
Matches 177; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

FT WO9418318-A. comments)"

PN 18-AUG-1994.

PD 01-FEB-1994; U00977.

PF 01-FEB-1993; US-013416.

PR 30-DEC-1993; US-176500.

PR 31-JAN-1994; US-189331.

PA (UYNC-) UNIV NORTH CAROLINA.

PI Fowlkes DM, Kay BK;

DR WPI; 94-279739/34.

PT Identifying proteins or peptide(s) which bind a ligand - by

PT screening a recombinant vector library expressing fusion proteins

PT comprising a binding domain and an effector domain

PS Disclosure; Page 35; 255pp; English.

CC Q70469 is a generic DNA sequence used to generate random TSAR peptide

CC This generic formula can be represented as follows: X(TGC)(NNB)10-

CC (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction

CC sites (X is not the same as Y) that are not specified further. This

CC sequence generates peptides that are cloverleaf in structure. Other

CC generic sequences are shown in Q70465-68. Other specific peptides

CC generated by these generic sequences are shown in R65150-54. TSARs are

CC concatenated heterofunctional proteins or peptides, comprising at least

CC two functional regions - a binding domain with affinity for a ligand and

CC a second effector peptide portion that is chemically or biologically

CC active.They may further comprise a linker peptide between the 2 domains.

CC The oligonucleotides are also designed so that the expressed peptide

CC contains 2 or 4 cysteine residues positioned in, or flanking, the

CC unpredicted or variant residues. These residues confer some degree of

CC conformational rigidity to the peptides. The TSARs or compsns. comprising

CC a TSAR binding domain can be used in vivo to deliver a chemically or

CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin

CC the function of macromolecules, eg. monoclonal or polyclonal antibodies

CC and therefore circumvent the need for complex methods of hybridoma

CC formation or in vivo antibody production. The TSARs are easily

CC characterised and have designed activity allowing direct and rapid

CC detection in a screening process.

SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 1.9%; Score 39; DB 12; Length 114;

Best Local Similarity 7.2%; Pred. No. 2.65e-06;

Matches 8; Conservative 32; Mismatches 71; Indels 0; Gaps 0;

Db 1 tgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 60

QY 595 TGCCACTCGCTCTGCGTGACGTTGAGGACCGCTGCGGCCCGGTGATGTCGCCCTTCGGC 654

Db 61 nnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 111

QY 655 TTCCCTGGCCCGACATGCTTGAGTGGACCGGTTTCCCCCAGGACAACGAC 705

RESULT 13

ID Q70467 standard; DNA; 114 BP.

AC Q70467;

DT 05-APR-1995 (first entry)

DE Generic DNA sequence to generate a random TSAR petide library.

KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;

KW effector domain; concatenated heterofunctional protein; linker;

KW direct; rapid; detection; screening; treatment; generic; ss.

OS Synthetic.

FH Key Location/Qualifiers

FT misc_feature 55..60

FT /*tag= a

FT /note= "this sequence represents 'Z'; Z can be a

FT sequence of 6, 9 or 12 nucleotides (see

FT comments)"

PN WO9418318-A.

PD 18-AUG-1994.

PF 01-FEB-1994; U00977.

PR 01-FEB-1993; US-013416.

PR 30-DEC-1993; US-176500.

PR 31-JAN-1994; US-189331.

PA (UYNC-) UNIV NORTH CAROLINA.

PI Fowlkes DM, Kay BK;

DR WPI; 94-279739/34.

DR P-PSDB; R65153.

PT Identifying proteins or peptide(s) which bind a ligand - by

PT screening a recombinant vector library expressing fusion proteins

PT comprising a binding domain and an effector domain

PS Disclosure; Page 35; 255pp; English.

CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally

CC Synthetic Affinity Reagents) peptides.This generic formula can also be

CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)1Y. X

CC and Y are flanking restriction sites (X is not the same as Y) that are

CC not specified further. Other generic sequences are shown in Q70466-68.

CC Other specific peptides generated by these generic sequences are shown in

CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,

CC comprising at least two functional regions - a binding domain with

CC affinity for a ligand and a second effector peptide portion that is

CC chemically or biologically active.They may further comprise a linker

CC peptide between the 2 domains.The oligonucleotides are also designed so

CC that the expressed peptide contains 2 or 4 cysteine residues positioned

CC in, or flanking, the unpredicted or variant residues. These residues

CC confer some degree of conformational rigidity to the peptides. The TSARs

CC or compsns. comprising a TSAR binding domain can be used in vivo to

CC deliver a chemically or biologically active moiety, eg. metal ion,

CC radioisotope, peptide, toxin or enzyme, to the specific target or on the

CC cell. They can also replace the function of macromolecules, eg.

CC monoclonal or polyclonal antibodies and therefore circumvent the need for

CC complex methods of hybridoma formation or in vivo antibody production.

CC The TSARs are easily characterised and have designed activity allowing

CC direct and rapid detection in a screening process.

SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 38; DB 12; Length 114;

Best Local Similarity 4.5%; Pred. No. 7.95e-06;

Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62

QY 198 GCCGCTGCGCCTTCCCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTGCGGCCACCAT 257

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114

QY 258 GCTGCAGGGCCCTGGCTGCTGCTGCTGCTCTTCTTCTGCTGCGCTGCGACTGCTGC 309

RESULT 14

ID Q70465 standard; DNA; 114 BP.

AC Q70465;

DT 05-APR-1995 (first entry)

DE Generic DNA sequence to generate a random TSAR petide library.

KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;

KW effector domain; concatenated heterofunctional protein; linker;

KW direct; rapid; detection; screening; treatment; generic; ss.

OS Synthetic.

FH Key Location/Qualifiers

FT misc_feature 55..60

FT /*tag= a

FT /note= "this sequence represents 'Z'; Z can be a

FT sequence of 6, 9 or 12 nucleotides (see

FT comments)"

PN WO9418318-A.

PD 18-AUG-1994.

PF 01-FEB-1994; U00977.

PR 01-FEB-1993; US-013416.

PR 30-DEC-1993; US-176500.

PR 31-JAN-1994; US-189331.

PA (UYNC-) UNIV NORTH CAROLINA.

PI Fowlkes DM, Kay BK;

DR WPI; 94-279739/34.

DR P-PSDB; R65150 and R65151.

PT Identifying proteins or peptide(s) which bind a ligand - by

PT screening a recombinant vector library expressing fusion proteins

PT comprising a binding domain and an effector domain

PS Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 7.95e-06;
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbntgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
QY 273 CTCGCTGCTGCTCTCTCTCGCCTCGCACCTGCTGCCTGGGCTCGGCGGGCTCTT 332

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbntgcnnbnnbnnb 114

QY 333 CCTCTTGGCCGCCGCGACTTCTCTCTACAGCGCAGCAATTGCAAGCCCATC 384

RESULT 15

ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are

CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or compsns. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 1.9%; Score 39; DB 12; Length 114;
Best Local Similarity 7.3%; Pred. No. 2.65e-06;
Matches 8; Conservative 32; Mismatches 69; Indels 0; Gaps 0;

Db 6 bnnbnnbnnbnnbnnbnnbnnbntgcnnbnnbnnbnnbnnbnnbnnbnnnn 65
Cp 540 CAGGAACCTTCTTGGTGTCGGGTGGCACCTGCTTCATGACGCGGATCCAAGCGCCGC 481

Db 66 btgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbntgc 114

Cp 480 CTGCTCCAGCACCTCCTTCATGGTCTCGTGGCCCGCAGCAGGTGGGCAGC 432

RESULT 16

ID Q70468 standard; DNA; 114 BP.
AC Q70468;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues

CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 7.95e-06;
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62

Cp 543 GCACAGGAACCTTCTTGGTGCCGGGTGGCACTGCTTCATGACCGGGATCCAGCGCC 484

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114

Cp 483 GGCCTGCTCCAGCACCTCTTCATGGTCTCGTGGCCCGCCAGCAGGTTGGCAGC 432

RESULT 17
ID Q70466 standard; DNA; 114 BP.
AC Q70466;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR-9 petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"

PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.

CC Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)1(TGC)(NNB)10(TGC)2(NNB)4Z(NNB)8(TGC)(NNB)
CC -9Y. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
CC Q70466-68. Other specific peptides generated by these generic sequences
CC are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active. They may further comprise a
CC linker peptide between the 2 domains. The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or compsns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.

CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 8.9%; Pred. No. 7.95e-06;
Matches 10; Conservative 31; Mismatches 71; Indels 0; Gaps 0;

Db 3 btgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 52

Cp 513 CTGCTTCATGACCAAGCGGGATCCAAAGCGCGCTGCTCCAGCACCTCTTCATGGTCTC 454

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114

Cp 453 GTGGCCCGCAGCAGGTTGGCGACGCCGATGTTCTGGTATTCCATGCGCGTGGCAC 402

RESULT 18
ID T76405 standard; DNA; 178 BP.
AC T76405;
DT 15-SEP-1997 (first entry)
DE Human endothelin-1 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; ss.
OS Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
DR WPI; 97-051871/05.

PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Claim 5; Page 38; 71pp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human endothelin-1, targeted at the initiation codon. The
CC method can be used to treat airway diseases such as cystic fibrosis,
CC asthma, chronic obstructive pulmonary disease, bronchitis and other
CC airway diseases characterised by an inflammatory response. By
CC eliminating adenosine from the antisense ON, its liberation upon
CC antisense degradation is prevented, thereby preventing adenosine-
CC induced bronchoconstriction in patients with hyper-reactive airways.
SQ Sequence 178 BP; 0 A; 52 C; 46 G; 32 T;

Query Match 1.9%; Score 39; DB 32; Length 178;
Best Local Similarity 33.3%; Pred. No. 2.65e-06;
Matches 20; Conservative 32; Mismatches 7; Indels 1; Gaps 1;

Db 113 bbbcbgcbgbbbtcbtgbgcbbbbtccbtctgbbbbbgbggbtcbbbbcc 172

QY 193 CTCGGCGCCTCGCCCTTCCCGGCTCCGCTCCCT-CTGCCCTCGGGGTGCGCGCC 251

RESULT 19
ID T76452 standard; DNA; 190 BP.
AC T76452;
DT 16-SEP-1997 (first entry)
DE Chymase antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; ss.
OS Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
DR WPI; 97-051871/05.

PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Example 5; Page 40; 71pp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for chymase, targeted at the initiation codon. The method can be
CC used to treat airway diseases such as cystic fibrosis, asthma, chronic
CC obstructive pulmonary disease, bronchitis and other airway diseases
CC characterised by an inflammatory response. By eliminating adenosine
CC from the antisense ON, its liberation upon antisense degradation is
CC prevented, thereby preventing adenosine-induced bronchoconstriction
CC in patients with hyper-reactive airways.
SQ Sequence 190 BP; 1 A; 41 C; 58 G; 36 T;

Query Match 1.9%; Score 39; DB 32; Length 190;
Best Local Similarity 40.0%; Pred. No. 2.65e-06;
Matches 36; Conservative 31; Mismatches 22; Indels 1; Gaps 1;

Db 28 bbgbbccctgbtcbtcbbcbgcttcgctggbgcbcbgbbgbbgcbgcbggbg 87
QY 208 CCTTCCCGGCTCCGCTCCCTCTGCCCC-CTCGGGTGGCGGCCACGATGTCAGGG 266
Db 88 gbgbgbbgcbgcbtcttcccbgbgbggc 117
QY 267 CCTGGCTCGCTGCTGCTTCTTCTCTCGC 296

RESULT 20
ID Q70470 standard; DNA; 114 BP.
AC Q70470;
DT 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "encoded by z (see comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNCL) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R58378.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 36; 255pp; English.
CC Q70470 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)4(CAC)(NNB)4(CAC)(NNB)8Z(NNB)6(CAC)(NNB)8
CC -(CAC)2(NNB)Y. X and Y are flanking restriction sites (X is not the same
CC as Y) that are not specified further. The peptides generated by this and
CC other generic sequences (Q70471-73) have invariant histidine residues
CC incorporated into variant sequences. TSARs are concatenated
CC heterofunctional proteins or peptides, comprising at least two functional
CC regions - a binding domain with affinity for a ligand and a second
CC effector peptide portion that is chemically or biologically active.They
CC may further comprise a linker peptide between the 2 domains. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need

CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 5 A; 10 C; 0 G; 0 T;

Query Match 1.8%; Score 36; DB 12; Length 114;
Best Local Similarity 10.2%; Pred. No. 6.96e-05;
Matches 11; Conservative 29; Mismatches 68; Indels 0; Gaps 0;

Db 3 bnnbnnbncacnbnbnbnbnbcacnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 62
QY 588 CCAGCCATGCCACTCGCTCTGCGTGAGGACCGCTGCGCCCGGTCATGTCGC 647
Db 63 bnnbnnbnnbnbnbnbcacnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbcacca 110
QY 648 CTTGGGCTTCCCTGGCCGACATGCTTGAGTGCGACCGGTTTCCCCCA 695

RESULT 21
ID Q70472 standard; DNA; 114 BP.
AC Q70472;
DT 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "encoded by z"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNCL) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R58383.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 36; 255pp; English.
CC Q70472 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)1(CAC)(NNB)11(CAC)(NNB)(CAC)(NNB)22(NNB)5
CC -(CAC)(NNB)5(CAC)2(NNB)4. X and Y are flanking restriction sites
CC (X is not the same as Y) that are not specified further. The peptides
CC generated by this and other generic sequences (Q70470-73) have invariant
CC histidine residues incorporated into variant sequences. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active.They may further comprise a linker peptide between the 2 domains.
CC The TSARs or compsns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 6 A; 12 C; 0 G; 0 T;

Query Match 1.8%; Score 36; DB 12; Length 114;
Best Local Similarity 10.9%; Pred. No. 6.96e-05;
Matches 12; Conservative 29; Mismatches 69; Indels 0; Gaps 0;

Db 5 acnnbnnbn 64
QY 623 ACCGCTGGCCCCGGTGCATGTCGGCTTCCCTGGCCGACATGCTTGAGTGC 682


```
FT WO9418318-A.
PN 18-AUG-1994.
PD 01-FEB-1994; U00977.
PE 01-FEB-1993; US-013416.
PR 01-FEB-1993; US-176500.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R58379.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 36; 255pp; English.
CC Q70471 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(CAC)(NNB)9(CAC)(NNB)Z(CAC)(NNB)4(CAC)2
CC -(NNB)6(CAC)(NNB)(CAC)(NNB)2Y. X and Y are flanking restriction sites
CC (X is not the same as Y) that are not specified further. The peptides
CC generated by this and other generic sequences (Q70470-73) have invariant
CC histidine residues incorporated into variant sequences. TSARS are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active.They may further comprise a linker peptide between the 2 domains.
CC The TSARS or compsns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARS are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 7 A; 14 C; 0 G; 0 T;

Query Match 1.7%; Score 34; DB 12; Length 114;
Best Local Similarity 12.5%; Pred. No. 5.87e-04;
Matches 14; Conservative 28; Mismatches 70; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbcaacnnbnnbnnbnnbnnbnnbnnbnnbcaacnnbnnnnnca 62
QY 348 CGACTTCTCCTACAGCGCAGCAATTGCAAGCCCATCCCGGCCAACCTGCAGTGTGCCA 407
Db 63 cnnbnnbnnbnnbcaacacnnbnnbnnbnnbnnbnnbnnbcaacnnbcaacnnbnnb 114
QY 408 CGGCATCGAATACCAAGCAATCGGGCTGCCCAACCTGCTGGGCCACGAGACC 459

RESULT 25
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PE 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.

PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 36; 255pp; English.
CC Q70471 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(CAC)(NNB)9(CAC)(NNB)Z(CAC)(NNB)4(CAC)2
CC -(NNB)6(CAC)(NNB)(CAC)(NNB)2Y. X and Y are flanking restriction sites
CC (X is not the same as Y) that are not specified further. The peptides
CC generated by this and other generic sequences (Q70470-73) have invariant
CC histidine residues incorporated into variant sequences. TSARS are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active.They may further comprise a linker peptide between the 2 domains.
CC The TSARS or compsns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARS are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 7 A; 14 C; 0 G; 0 T;
```

```
DR P-PSDB; R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARS are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARS
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARS are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.7%; Score 34; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 5.87e-04;
Matches 4; Conservative 33; Mismatches 75; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 62
Cp 630 GCAGCGGTCTTCACCTGCAGCGCAGAGCGAGTGGCATGGTGATGGTCTCGTCTAGGTC 571
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
Cp 570 ATCGAGGCAGACGGGGCGGAGAGCGGAGCACAGGAACTCTTGGTGTCCGGG 519

RESULT 26
ID Q70471 standard; DNA; 114 BP.
AC Q70471;
DT 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "encoded by 2 (see comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PE 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R58379.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 36; 255pp; English.
CC Q70471 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(CAC)(NNB)9(CAC)(NNB)Z(CAC)(NNB)4(CAC)2
CC -(NNB)6(CAC)(NNB)(CAC)(NNB)2Y. X and Y are flanking restriction sites
CC
```



```
FT      /*tag= e
FT      /label= mRNA start sequence
FT      /note= "501 is possible start site"
FT      504..789
FT      /*tag= f
FT      /label= mRNA start sequence
FT      /note= "504 is possible start site"
FT      506..789
FT      /*tag= g
FT      /label= mRNA start sequence
FT      /note= "506 is possible start site"
FT      790..3498
FT      /*tag= h
FT      /label= HSV-1 gB
FT      /note= "includes N-terminal hydrophobic leader and
FT      a membrane-spanning sequence, a C-terminal
FT      ionic sequence, and 9 N-linked
FT      saccharide-addition sites"
FT      3499..3549
FT      /*tag= i
FT      /label= poly_a_signal
FT      3518..3525
FT      /*tag= j
FT      /label= poly_a_site
FT      3549..3549
FT      /*tag= k
FT      /label= misc_rna
FT      3549..3997
FT      /*tag= l
FT      /note= "3' nonessential sequences to the BamHI
FT      site"
PN      US4642333-A.
PD      10-FEB-1987.
PF      20-JUN-1984; 622496.
PR      16-SEP-1983; US-532996.
PR      20-JUN-1985; US-622496.
PA      (PERS/) PERSON S.
PI      Person S;
DR      WPI; 87-056354/08.
PT      Amino acid chain of glyco:protein B of HSV-1 and 2 - prepd. as
PT      recombinant and used for vaccines for herpes simplex virus types 1
PT      and 2.
PS      Example; Table 1; 16pp; English.
CC      !NOTE! This sequence has been indexed as represented in the
CC      specification, except that bases 'E' have been replaced by 'N'.
CC      The features have been indexed according to the legend of table 1 on
CC      column 19/20 and the Sequence Summary of column 5 (sic). Note that
CC      the base numbering of the features does not correspond to the
CC      the sequence numbering below.
CC      For another DNA sequence of HSB-1 gB see N71303 (P71135),
CC      and for HSV-2 gB see N71399 (P71136).
CC      A pure non-glycosylated amino acid (AA) chain comprising a sequence
CC      corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
CC      virus which is antigenic to HSV-1 of HSV-2, which contains no more
CC      than 750 AA residues, and which includes AA residues 135-649
CC      inclusive is claimed. It can be used to produce vaccines for
CC      prophylaxis and treatment of HSV-1 and HSV-2.
SQ      Sequence 3871 BP; 743 A; 1402 C; 544 G; 754 T;

Query Match 1.5%; Score 30; DB 2; Length 3871;
Best Local Similarity 37.3%; Pred. No. 3.65e-02;
Matches 47; Conservative 31; Mismatches 48; Indels 0; Gaps 0;

Db 2781 bgaggtccacbccbcaaccabtbacbacbcbttcbaccatcacatcacgtcatcc 2840
:| | | | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 189 CGAACTCGGCGCGCTCGCCCTTCCCGGCTCCGCTCCGCTCCCTCTGCCCTCGGGTGGCGC 248

Db 2841 acbcgacbccacbcccatsttcbbcbcbcbcbcbcbcbcbcbcbcbcbcbcbcb 2900
|:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 249 GCCCAGCATGCTGCAGGGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2900
|:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 2901 cctbgg 2906
|:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 309 CCTGGG 314
```

```
RESULT 39
ID T76167 standard; DNA; 69 BP.
AC T76167;
DT 12-SEP-1997 (first entry)
DE Human IL3 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; interleukin; ss.
OS Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
DR WPI; 97-051871/05.
PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Claim 5; Page 28; 7lpp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human IL3, targeted at the initiation codon. The method can be
CC used to treat airway diseases such as cystic fibrosis, asthma, chronic
CC obstructive pulmonary disease, bronchitis and other airway diseases
CC characterised by an inflammatory response. By eliminating adenosine from
CC the antisense ON, its liberating adenosine-induced bronchoconstriction is
CC prevented, thereby preventing adenosine-induced bronchoconstriction in
CC patients with hyper-reactive airways.
SQ Sequence 69 BP; 0 A; 15 C; 31 G; 9 T;

Query Match 1.4%; Score 28; DB 32; Length 69;
Best Local Similarity 50.0%; Pred. No. 2.67e-01;
Matches 31; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

Db 6 bcbggbgttggbgcbggbgcbggbgcbggbgcbggbgcbggbgcbggbgcbggbgcbg 65
:| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 295 GCCTCGCACTGCTGCCTGGGCTCGGCGCGGGGCTCTTCTCTTGGCCAGCCGACTTC 354

Db 66 bc 67
:|
QY 355 TC 356

RESULT 40
ID T76184 standard; DNA; 70 BP.
AC T76184;
DT 12-SEP-1997 (first entry)
DE Human IL4 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; interleukin; ss.
OS Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
DR WPI; 97-051871/05.
PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Claim 5; Page 29; 7lpp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human IL4, targeted at the initiation codon. The method can be
CC used to treat airway diseases such as cystic fibrosis, asthma, chronic
CC obstructive pulmonary disease, bronchitis and other airway diseases
CC characterised by an inflammatory response. By eliminating adenosine from
CC the antisense ON, its liberation upon antisense degradation is
```



```

||||| ||||| ||| | || |||| ||||| ||||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTAAAA 1839

RESULT 50
ID N80801 standard; DNA; 1466 BP.
AC N80801;
DE 12-SEP-1990 (first entry)
DT Sequence encoding vascular anticoagulating protein (VAP) VAC-alpha
DE (Glu22)
KW Vascular anticoagulating protein (VAP) VAC-alpha; anticoagulant;
KW antiinflammatory; phospholipase inhibitory.
OS Homo sapiens.
FH Key Location/Qualifiers
FT mat_peptide 35..997
FT /*tag= a
FT /note="claimed in claim 2, page 160-161"
FN WO8807576-A.
PD 06-OCT-1988.
PF 26-MAR-1988; E00266.
PR 04-NOV-1987; DE-737367,
PA (BOEH) Boehringer Ingelheimint.
PI Hauptmann R, Maurer-Fogy I, Bodo G, Swetly P, Stratowa C, Falkner E,
PI Adolf G, Reutlingsperger CMP;
DR WPI: 88-292861/41.
DR P-PSDB; P80714.
PT New vascular anti-coagulating proteins -
PT useful as thrombin inhibitors, antiinflammatory agents, etc
PS Disclosure; Fig 4/1-4/3; 183pp; German.
CC The DNA is produced by determining the amino acid sequence of isolated
CC and highly purified VAPs, constructing DNA probes on the basis of this
CC sequence, using the probes to search through suitable cDNA libraries,
CC isolating cDNA that hybridises with the probes, by inserting the cDNA
CC into a suitable vector, and using the vector to transform a host
CC organism. VAC-alpha has anticoagulant activity under certain
CC conditions, but loses this activity in the event of severe bleeding.
CC It acts by inhibiting conversion of factor X to factor Xa and conversion
CC of prothrombin to thrombin. It is structurally related to lipocortin I
CC (Nature, 320, 77, 1986) and lipocortin II (Cell, 46, 191/120, 1986) and
CC has similar antiinflammatory and phospholipase inhibitory activity.
SQ Sequence 1466 BP; 422 A; 289 C; 328 G; 424 T; 3 Others;

Query Match 1.4%; Score 28; DB 1; Length 1466;
Best Local Similarity 80.4%; Pred. No. 2.67e-01;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 1252 atatatatttttaagagcctagagtgccttttagccttttttaaaa 1297
||||| ||||| ||| | || |||| ||||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTAAAA 1839

RESULT 51
ID N81113 standard; DNA; 1567 BP.
AC N81113;
DT 12-NOV-1990 (first entry)
DE Plasmid contg. placental coagulation inhibitor gene.
KW Placental coagulation inhibitor; disseminated vascular coagulation;
KW thrombosis; ss.
FH Key Location/Qualifiers
FT cds 136..1098
FT /*tag= a
FT /label=placental coagulation inhibitor
FT 1568
FT /*tag= b
FN EP-279459-A.
PN J03219875-A.
PN J03219875-A.
PD 24-AUG-1988.
PF 19-FEB-1988; 102468.
PR 23-JUL-1987; JP-184428.
PR 23-JUL-1987; JP-037227.
PA (KOWA) Kowa KK.
PI Saino Y, Iwasaki A, Suda M;
DR WPI: 88-236733/34.

```

```

WPI; 91-329110/45.
DR P-PSDB; P80511.
PT Recombinant placental coagulation inhibitor - useful for the prevention
PT and treatment of thromboses or disseminated intra-vascular coagulation.
PS Disclosure; p; English.
CC This recombinant DNA sequence comprises a promoter region, ribosome
CC binding site, initiation sequence, termination codon and a trans-
CC cription terminator in addition to the placental coagulation inhibitor
CC gene. The polypeptide produced on culturing of transformed host cells
CC exhibits strong anticoagulant activities and is useful for the treat-
CC ment and prevention of eg thrombosis or disseminated intravascular
CC coagulation in the brain, heart and peripheral blood vessels.
SQ Sequence 1567 BP; 430 A; 332 C; 358 G; 447 T;

Query Match 1.4%; Score 28; DB 1; Length 1567;
Best Local Similarity 80.4%; Pred. No. 2.67e-01;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 1353 atatatatttttaagagcctagagtgccttttagccttttttaaaa 1398
||||| ||||| ||| | || |||| ||||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTAAAA 1839

RESULT 52
ID N90112 standard; DNA; 1575 BP.
AC N90112;
DT 1-NOV-1989 (first entry)
DE Anticoagulant PP4 DNA
KW Anticoagulant; PP4 protein; thromboplastin.
KW Homo sapiens (human)
OS Location/Qualifiers
FH Key 1539..1575
FT polya_signal /*tag= a
FT cds 107..1066
FT /*tag= b
FN EP-318703-A.
PD 07-JUN-1989.
PF 29-SEP-1988; 118039.
PR 03-NOV-1987; EP-737239.
PA (BEHW) Behringwerke AG.
PI Grundmann U, Abel K-J, Kupper H;
DR WPI; 89-166767/23.
PT New DNA sequence encoding anticoagulant PP4 protein
PT - and new recombinant protein, vectors, antibodies, etc.,
PT useful therapeutically and diagnostically.
PS Claim 1; table 1; 14pp; German.
CC DNA sequence encoding anticoagulant PP4 protein. This
CC inhibits blood coagulation at the thromboplastin stage.
CC The derived protein is 320 amino acids.
SQ Sequence 1575 BP; 461 A; 322 C; 351 G; 441 T;

Query Match 1.4%; Score 28; DB 1; Length 1575;
Best Local Similarity 80.4%; Pred. No. 2.67e-01;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 1324 atatatatttttaagagcctagagtgccttttagccttttttaaaa 1369
||||| ||||| ||| | || |||| ||||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTAAAA 1839

RESULT 53
ID N91821 standard; DNA; 1605 BP.
AC N91821;
DT 16-MAR-1990 (first entry)
DE Endonexin II complete cDNA
KW Placenta; blood coagulation.
OS Homo sapiens (human).
FH Key Location/Qualifiers
FT cds 160..1119
FT /*tag= a
FN EP-339285-A.
PD 02-NOV-1989.

```

```
PF 30-MAR-1989; 105626.
PR 31-MAR-1988; US-176802.
PA (RORE) Rorer Int Overseas.
PI Kaplan R, Jaye M;
DR WPI; 89-317181/44.
DR P-PSDB; P91021.
PT Recombinant human endonexin II - which inhibits blood coagulation
PT and phospholipase A2 and is used as an antiinflammatory agent
PS Claim 1; fig. 1; 13pp; English.
CC hENII is a calcium and phospholipid binding protein.
SQ Sequence 1605 BP; 447 A; 337 C; 366 G; 455 T;

Query Match 1.4%; Score 28; DB 1; Length 1605;
Best Local Similarity 80.4%; Pred. No. 2.67e-01;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 1378 atatatattttaagagcgtagagtgtttagccttttttaaaa 1423
    ||||| ||||| ||| | || |||| ||||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGAGTACTTAAGCCTTGTTTAAA 1839
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 54
ID Q65611 standard; cDNA; 1840 BP.
AC Q65611;
DT 01-FEB-1995 (first entry)
DE Feline zona pellucida ZPB coding sequence.
KW Cat; feline; zona pellucida; ZPB; immunocontraception; ds.
OS Felis domesticus.
FH Key Location/Qualifiers
FT cds 57..1769
FT /*tag= a
FT /product= ZPB
PN WO9411019-A.
PD 26-MAY-1994.
PF 06-NOV-1993; U10851.
PR 09-NOV-1992; US-973341.
PR 29-JAN-1993; US-012990.
PA (ZONA-) ZONAGEN INC.
PI Harris JD, Hsu KT, Podolski JS;
DR WPI; 94-183156/22.
DR P-PSDB; R55201.
PT Use of zona pellucida proteins and antibodies - for inducing
PT reproducible transient infertility or permanent sterility in
PT female mammals
PT Claim 24; Page 99-101; 154pp; English.
PS A cDNA library was prepared in lambda gt10 from mRNA isolated from
CC ovaries of 3-4 month old cats. Plaques were screened using a
CC mixture of probes encoding porcine ZPA, ZPB and ZPC proteins.
CC Positive clones were analysed further by Southern hybridisation
CC using the porcine probes and clones encoding feline ZPA, ZPB and
CC ZPC proteins were identified. The deduced amino acid sequence from
CC the feline ZPB clone was approximately 80% homologous to porcine ZPB
CC protein.
SQ Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T;

Query Match 1.4%; Score 29; DB 11; Length 1840;
Best Local Similarity 96.8%; Pred. No. 9.94e-02;
Matches 30; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1804 ttaccaaaaaaaaaaaaaaaaaaagggccgc 1834
    ||||| ||||| ||||| ||||| ||||| |||||
QY 1997 TTACCTAAAAAATAAAAAAAGGGCCGC 2027

RESULT 55
ID T89891 standard; DNA; 2259 BP.
AC T89891;
DT 27-APR-1998 (first entry)
DE Mouse frizzled gene 7 (Mfz7) encoding a Wnt receptor.
KW Wnt receptor; mouse frizzled gene 7; Mfz7 gene;
KW signal transduction; cancer; cell growth; cell proliferation;
KW mammary tumour; oncogene; therapy; ds.
OS Mus musculus.
```

```
FH Key Location/Qualifiers
FT CDS 361..2079
FT /*tag= a
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO ) UNIV JOHNS HOPKINS.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI; 97-526631/48.
DR P-PSDB; W31273.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 43-44; 61pp; English.
CC This sequence comprises novel mouse frizzled gene 7 (Mfz7)
CC that encodes a transmembrane receptor, frizzled-7 (see W31273),
CC a Wnt receptor (Wntr). Novel frizzled family members have been
CC identified in Drosophila, mouse, human and Caenorhabditis (see
CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
CC can be used in a novel, claimed method of screening for compounds
CC which modulate the binding of a Wnt polypeptide (secreted proteins
CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
CC involved in (mammary) cancer and other processes involving growth,
CC development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 2259 BP; 362 A; 739 C; 677 G; 481 T;

Query Match 1.4%; Score 29; DB 37; Length 2259;
Best Local Similarity 59.3%; Pred. No. 9.94e-02;
Matches 140; Conservative 0; Mismatches 93; Indels 3; Gaps 2;

Db 559 ctgcccacctgctggccacacagagagcgggcctcgaggtgcaccagttc 618
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 433 CTGCCCAACCTGCTGGCCACGAGACCATGAAGGAGTGCTGGAGCGCGGCTGG 492
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 619 taccctctggtaaagtgagtgagttctcctgagctacgcttcttctatgctctatgtac 678
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 493 ATCCCGCTGGTCATGAGCAGTGCCACCCGGACACCAAGAAGTTCTGTGCTCGCTCTC 552
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 679 gcacccgctgtgcaccg-tg--ctcgaccaagccattcctccgtgcccgttccttgcgag 735
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 553 GCCCCCGTCTGCCTCGATGACCTAGACGAGACCAATCCAGCCATGCCACTCGCTCTGCGTG 612
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 736 cgcgcccgcacagggtgcgagggcgctcatgaacaagttcggcttcaggggccaga 791
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 613 CAGGTGAAGGACCGCTGCGCCCGGTCATGTCCGCTTGGCTTCCCTTGCCCGCA 668
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 56
ID T76423 standard; DNA; 125 BP.
AC T76423;
DT 16-SEP-1997 (first entry)
DE Human endothelin ETA receptor antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; ss.
OS Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
DR WPI; 97-051871/05.
PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Example 5; Page 39; 71pp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
```

 W E S R E (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
 Run on: Thu Nov 5 10:57:48 1998; MasPar time 2269.27 Seconds
 1370.520 Million cell updates/sec
 Tabular output not generated.

Title: >US-08-848-439-1
 Description: (1-2027) from US08848439.seq
 Perfect Score: 2027
 N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGCGGCCGC 2027
 Comp: CTTAAGCCGGAAGTACCGGA.....TTTTTTTTTTTCGCCGGCG

Scoring table: TABLE default
 Gap 6

Nmatch STD : Dbase 0; Query 0
 Searched: 1988617 seqs, 767163441 bases x 2
 Post-processing: Minimum Match 0%
 Listing first 60 summaries

Database: embl-est55
 Database: 1:em_est10 2:em_est11
 3:gb_est1 4:gb_est10 5:gb_est11 6:gb_est12 7:gb_est13
 8:gb_est14 9:gb_est15 10:gb_est16 11:gb_est17
 12:gb_est18 13:gb_est19 14:gb_est2 15:gb_est20
 16:gb_est21 17:gb_est22 18:gb_est23 19:gb_est24
 20:gb_est3 21:gb_est4 22:gb_est5 23:gb_est6 24:gb_est7
 25:gb_est8 26:gb_est9 27:gb_gss

Statistics: Mean 12.290; Variance 3.203; scale 3.837
 Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	553	27.3	579 7	AA528226 nh92a05.s1 NCI_CGAP_Br
2	479	23.6	505 25	AA071469 ze88f06.s1 Soares feta
3	473	23.3	480 22	N93894 yy64d09.r1 Homo sapien
4	472	23.3	505 24	W92531 ze03g10.r1 Soares feta
5	468	23.1	473 8	AA513750 nh88g03.s1 NCI_CGAP_Br
6	468	23.1	474 15	AA857092 of62a09.s1 NCI_CGAP_Co
7	453	22.3	460 24	W72225 zd70e11.s1 Soares feta
8	434	21.4	443 7	AA449300 zx08c06.s1 Soares tota
9	433	21.4	482 17	AA922948 chl0h11.s1 NCI_CGAP_Co
10	425	21.0	427 5	AA431859 zw51g08.s1 Soares tota
11	420	20.7	518 18	AA927991 on47g08.s1 NCI_CGAP_Co
12	416	20.5	446 22	N64004 yy64d09.s1 Homo sapien
13	413	20.4	462 24	W77968 zd70e11.r1 Soares feta

C	14	409	20.2	409	7	AA449749	zx07e10.s1 Soares tota	0.00e+00
C	15	408	20.1	445	21	N32415	yw83c06.s1 Homo sapien	0.00e+00
C	16	407	20.1	562	7	AA449032	zx07e10.r1 Soares tota	0.00e+00
C	17	405	20.0	405	8	AA505877	ni01a05.s1 NCI_CGAP_Br	0.00e+00
C	18	402	19.8	408	17	AA976403	oq68f09.s1 NCI_CGAP_Ki	0.00e+00
C	19	388	19.1	388	7	AA528219	nj16h09.s1 NCI_CGAP_Pr	0.00e+00
C	20	366	18.1	422	5	AA431858	zw51g08.r1 Soares tota	0.00e+00
C	21	366	18.1	458	25	AA105749	ml84a03.r1 Stratagene	0.00e+00
C	22	362	17.9	378	10	AA429960	zw58g10.r1 Soares tota	0.00e+00
C	23	363	17.9	433	23	AA024772	ze76g10.s1 Soares feta	0.00e+00
C	24	359	17.7	426	21	H87071	ys74g07.r1 Homo sapien	0.00e+00
C	25	350	17.3	359	4	AA295018	EST100474 Pancreas tum	0.00e+00
C	26	338	16.7	449	18	AI047549	uh80b08.r1 Soares mous	0.00e+00
C	27	337	16.6	406	20	H16121	yl28c05.r1 Homo sapien	0.00e+00
C	28	331	16.3	340	6	HUM296B07B	Human aorta CDNA 5'-en	0.00e+00
C	29	328	16.2	396	26	AA073862	mm97f07.r1 Stratagene	0.00e+00
C	30	329	16.2	414	23	AA024771	ze76g10.r1 Soares feta	0.00e+00
C	31	323	15.9	329	12	AA705737	zf41b02.s1 Soares feta	0.00e+00
C	32	322	15.9	434	20	H44092	yo73a07.r1 Homo sapien	0.00e+00
C	33	321	15.8	321	4	AA335712	EST40158 Epididymus Ho	0.00e+00
C	34	320	15.8	328	4	AA295688	EST100887 Pancreas tum	0.00e+00
C	35	321	15.8	469	20	H15818	yl28c05.s1 Homo sapien	0.00e+00
C	36	319	15.7	326	21	N32424	yw83e07.s1 Homo sapien	0.00e+00
C	37	316	15.6	330	6	HUM230F10B	Human aorta CDNA 5'-en	0.00e+00
C	38	313	15.4	344	22	N56835	yw83c06.r1 Homo sapien	0.00e+00
C	39	299	14.8	301	6	HUM303B01B	Human aorta CDNA 5'-en	0.00e+00
C	40	301	14.8	301	10	AA602994	np31g01.s1 NCI_CGAP_Pr	0.00e+00
C	41	299	14.8	313	6	HUM304B12B	Human aorta CDNA 5'-en	0.00e+00
C	42	298	14.7	303	6	HUM347H05B	Human aorta CDNA 5'-en	0.00e+00
C	43	296	14.6	300	4	AA347786	EST54420 Fetal heart I	0.00e+00
C	44	294	14.5	300	6	HUM307G04B	Human aorta CDNA 5'-en	0.00e+00
C	45	292	14.4	302	4	AA330758	EST34493 Embryo, 6 wee	0.00e+00
C	46	290	14.3	293	6	HUM238C04B	Human aorta CDNA 5'-en	0.00e+00
C	47	290	14.3	293	6	HUM238C05B	Human aorta CDNA 5'-en	0.00e+00
C	48	290	14.3	344	20	H14917	yl26d03.r1 Homo sapien	0.00e+00
C	49	287	14.2	299	6	HUM240B02B	Human aorta CDNA 5'-en	0.00e+00
C	50	285	14.1	287	6	HUM289F09B	Human aorta CDNA 5'-en	0.00e+00
C	51	279	13.8	281	6	HUM343G04B	Human aorta CDNA 5'-en	0.00e+00
C	52	273	13.5	273	6	HUM233G01B	Human aorta CDNA 5'-en	0.00e+00
C	53	274	13.5	276	5	AA371169	EST83193 Prostate glan	0.00e+00
C	54	273	13.5	282	4	AA339113	EST44185 Fetal brain I	0.00e+00
C	55	269	13.3	283	6	Cl6543	Human aorta CDNA 5'-en	0.00e+00
C	56	256	12.6	262	25	AA176165	zp23h12.s1 Stratagene	0.00e+00
C	57	256	12.6	263	6	HUM286A04B	Human aorta CDNA 5'-en	0.00e+00
C	58	255	12.6	287	25	AA082155	ze88f06.r1 Soares feta	0.00e+00
C	59	255	12.6	394	26	AA260087	va52g08.r1 Soares mous	0.00e+00
C	60	254	12.5	266	6	HUM224H07B	Human aorta CDNA 5'-en	0.00e+00

ALIGNMENTS

RESULT 1
 LOCUS AA528226 579 bp mRNA EST 05-AUG-1997
 DEFINITION nh92a05.s1 NCI_CGAP_Br.1 Homo sapiens cDNA clone IMAGE:965936,
 mRNA sequence.
 ACCESSION AA528226
 NID 92270295
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
 Homo.
 REFERENCE 1 (bases 1 to 579)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 682 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 389.

FEATURES
source
1. .505
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."
/db_xref="taxon:9606"
/clone="357954"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1. .>505

mRNA
BASE COUNT 149 a 95 c 81 g 180 t
ORIGIN

Query Match 23.3%; Score 472; DB 24; Length 505;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 498; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 ATTTTATGAAGTTTAAAAATAGCTCATTAAAGCTAGTTTTGAATAGGTGCAACTGTG 60
QY 1402 ATTTTATGAAGTTTAAAAATAGCTCATTAAAGCTAGTTTTGAATAGGTGCAACTGTG 1461
Db 61 ACTTGGGTCTGGTTGGTTGTTGTTGTTTGGTTCAGTCAGCTGATTTTCACITCCCACTG 120
QY 1462 ACTTGGGTCTGGTTGGTTGTTGTTGTTTGGTTCAGTCAGCTGATTTTCACITCCCACTG 1521
Db 121 AGTTGTCATAACATGCAAAATGCTTCAATTTTCTCTGTGGCCCCAAACTTGTGGGTCA 180
QY 1522 AGTTGTCATAACATGCAAAATGCTTCAATTTTCTCTGTGGCCCCAAACTTGTGGGTCA 1581
Db 181 AACCCCTGTTGAGATAAAGCTGGCTGTTATCTCAACATCTTCATCAGCTCCAGACTGAGAC 240
QY 1582 AACCCCTGTTGAGATAAAGCTGGCTGTTATCTCAACATCTTCATCAGCTCCAGACTGAGAC 1641
Db 241 TCAGTGTCTAAGTCTTACAACAATTCATCATTTTTATACCTTCAATGGGAACTTAAACTGT 300
QY 1642 TCAGTGTCTAAGTCTTACAACAATTCATCATTTTTATACCTTCAATGGGAACTTAAACTGT 1701
Db 301 TACATGTATCACATTCAGCTACAACACTACTTCCATTTATTAAAGAGCACATTAAACCATTT 360
QY 1702 TACATGTATCACATTCAGCTACAA-TACTTCCATTTATTA-GAAGCACATTAAACCATTT 1759
Db 361 CTATAGCATGATTTCTCAAGTAAAGGCAAGATATATAAAATTTTATAATTGACTTGAGT 420
QY 1760 CTATAGCATGATTTCTCAAGTAAAGGCAAGATATATAAAATTTTATAATTGACTTGAGT 1819
Db 421 ACTTTAAGCCTTGTTTAAACATTTCTTACTTTAACTTTTGCAAATTTAAACCCATTTGGTAG 480
QY 1820 ACTTTAAGCCTTGTTTAAACATTTCTTACTTTAACTTTTGCAAATTTAAACCCATTTG-TAG 1878
Db 481 CCTACCCGGTAATATACATAGTAG 504
QY 1879 CTTACCTG-TAATATACATAGTAG 1901

RESULT 5
LOCUS AA513750 473 bp mRNA EST 20-AUG-1997
DEFINITION nh88g03.s1 NCI_CGAP_Br1.1 Homo sapiens cDNA clone IMAGE:965620,

ACCESSION
NID
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 473)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 2969 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 472.
Location/Qualifiers
1. .473
/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is not normalized. (The normalized version of this library is NCI_CGAP_Br2.) Library was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:965620"
/clone_lib="NCI_CGAP_Br1.1"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
<1. .>473

mRNA
BASE COUNT 169 a 57 c 91 g 156 t
ORIGIN

Query Match 23.1%; Score 468; DB 8; Length 473;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGGTAAACACAGGATGTAAAGTTTATATACAAGAATAATGTTTATCTGAAATATTACA 60
Cp 2002 AGGTAAACACAGGATGTAAAGTTTATATACAAGAATAATGTTTATCTGAAATATTACA 1943
Db 61 GTGTTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGG 120
Cp 1942 GTGTTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGG 1883
Db 121 TAAGCTACAATGGGTTTAATTTGCAAAAGTTTAAGTAAGAAATGTTTTAAACAAGGCTTAA 180
Cp 1882 TAAGCTACAATGGGTTTAATTTGCAAAAGTTTAAGTAAGAAATGTTTTAAACAAGGCTTAA 1823
Db 181 AGTACTCAAGTCAATATAAAATTTATATCTTTTGCCTTTTACTTTGAAGAAATCATGCTA 240
Cp 1822 AGTACTCAAGTCAATATAAAATTTATATCTTTTGCCTTTTACTTTGAAGAAATCATGCTA 1763
Db 241 TAGAAATGGTTAATGTGCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGT 300

Cp 1762 TAGAAATGGTTAATGTGCTTCTAATAAATGGAAGTATTGTAGCTGGAAATGTGATACATGT 1703
|||||
Db 301 AACAGTTTAAGTTCCTCCATTTGAAGGTATATAAATGATGAATTTGTTGTAAGACTTAGACACTG 360
|||||
Cp 1702 AACAGTTTAAGTTCCTCCATTTGAAGGTATATAAATGATGAATTTGTTGTAAGACTTAGACACTG 1643
|||||
Db 361 AGTCTCCGCTCTGGAGCTGATGAAGATGTTGAGATAAACAGCCAGCTTTATCTCAACAGGGT 420
|||||
Cp 1642 AGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATAAACAGCCAGCTTTATCTCAACAGGGT 1583
|||||
Db 421 TTGTGACCCACAAAGTTTGGGCCCCCAGAGAAAAATTGAAGCAATTTGCGATGTTA 472
|||||
Cp 1582 TTGTGACCCACAAAGTTTGGGCCACAGAGAAAAATTGAAGCAATTTGCGATGTTA 1531
|||||
RESULT 6
LOCUS AA857092 474 bp mRNA EST 09-MAR-1998
DEFINITION Of62a09.s1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1434904 3',
mRNA sequence.
ACCESSION AA857092
NID g2945394
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 474)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 464.
Location/Qualifiers
1. .474
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1434904"
/clone_lib="NCI_CGAP_Co8"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
BASE COUNT 171 a 54 c 92 g 157 t
ORIGIN
Query Match 23.1%; Score 468; DB 15; Length 474;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 474; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 1 AGGTAAACAGGATGTAAAGTTTATATACAGAATAATAATGTTTATCTCGAATAATTACA 60
|||||
Cp 2002 AGGTAAACAGGATGTAAAGTTTATATACAGAATAATAATGTTTATCTCGAATAATTACA 1943
|||||

Db 61 GTGTTGGTTAAAGCAATATTTTTTACAACCTTTTAAAGGTTAACTACTATGTATATTACAGG 120
|||||
Cp 1942 GTGTTGGTTAAAGCAATATTTTTTACAACCTTTTAAAGGTTAACTACTATGTATATTACAGG 1883
|||||
Db 121 TAAGCTACAATGGGTTTAATTTGCAAAAAGTTAAGTAAGAAATGTTTTTAAACAAGGCTTAA 180
|||||
Cp 1882 TAAGCTACAATGGGTTTAATTTGCAAAAAGTTAAGTAAGAAATGTTTTTAAACAAGGCTTAA 1823
|||||
Db 181 AGTACTCAAGTCAATTATAAAATTTATATCTTTTGCCCTTTTACTTGAAGAAATCATGCTA 240
|||||
Cp 1822 AGTACTCAAGTCAATTATAAAATTTATATCTTTTGCCCTTTTACTTGAAGAAATCATGCTA 1763
|||||
Db 241 TAGAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGT 300
|||||
Cp 1762 TAGAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGT 1703
|||||
Db 301 AACAGTTTAAAGTTCCCATTTGAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACTG 360
|||||
Cp 1702 AACAGTTTAAAGTTCCCATTTGAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACTG 1643
|||||
Db 361 AGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATAAACAGCCAGCTTTTATCTCAACAGGGT 420
|||||
Cp 1642 AGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATAAACAGCCAGCTTTTATCTCAACAGGGT 1583
|||||
Db 421 TTGTGACC-ACAAAGTTTGGGCCACAGAGAAAAATTGAAGCAATTTGCGATGTTATGA 474
|||||
Cp 1582 TTGTGACCCACAAAGTTTGGGCCACAGAGAAAAATTGAAGCAATTTGCGATGTTATGA 1528
|||||
RESULT 7
LOCUS W72225 460 bp mRNA EST 17-OCT-1996
DEFINITION zd70e11.s1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
346028 3', mRNA sequence.
ACCESSION W72225
NID g1382674
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 384.
Location/Qualifiers
1. .460
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

/db_xref="taxon:9606"
/clone="346028"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
complement(<1.>460)

mRNA	164 a	55 c	87 g	154 t
BASE COUNT				
ORIGIN				
Query Match 22.3%; Score 453; DB 24; Length 460;				
Best Local Similarity 99.8%; Pred. No. 0.00e+00;				
Matches 459; Conservative 0; Mismatches 0; Indels 1; Gaps 1;				
Db	1	AAAGTTTATATACAAGAATAAATGTTTATCTGAAATATTTACAGTGTGGTTAAAGCAA	60	
Cp	1986	AAAGTTTATATACAAGAATAAATGTTTATCTGAAATATTTACAGTGTGGTTAAAGCAA	1927	
Db	61	TATTTTTCACAACTTTTAAAGGTAAACTACTATGTATATTCACAGGTAAGCTACAATGGGTT	120	
Cp	1926	TATTTTTCACAACTTTTAAAGGTAAACTACTATGTATATTCACAGGTAAGCTACAATGGGTT	1867	
Db	121	TAATTTGCAAAAGTTAAGTAAGAAATGTTTTTAAACAAGGTTAAAGTACTCAAGTCAATT	180	
Cp	1866	TAATTTGCAAAAGTTAAGTAAGAAATGTTTTTAAACAAGGTTAAAGTACTCAAGTCAATT	1807	
Db	181	ATAAAATTTATATCTTTTGCCCTTTTACTTTGAAGAAATCATGCTATAGAAATGGTTAATGT	240	
Cp	1806	ATAAAATTTATATCTTTTGCCCTTTTACTTTGAAGAAATCATGCTATAGAAATGGTTAATGT	1747	
Db	241	GCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAACAGTTTAAAGTCCC	300	
Cp	1746	GCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAACAGTTTAAAGTCCC	1687	
Db	301	ATTGAAGGTATAAAATGATGAATTCGTTGTAAGACTTAGACACTGAGTCTCAGTCTGGAG	360	
Cp	1686	ATTGAAGGTATAAAATGATGAATT-GTTGTAAGACTTAGACACTGAGTCTCAGTCTGGAG	1628	
Db	361	CTGATGAAGATGTTGAGATAACAGCCAGCTTTTATCTCAACAGGGTTTGTGACCCACAAGT	420	
Cp	1627	CTGATGAAGATGTTGAGATAACAGCCAGCTTTTATCTCAACAGGGTTTGTGACCCACAAGT	1568	
Db	421	TTGGGCCACAGAGAAAATTGAAGCAATTTCATGATGTTATGA	460	
Cp	1567	TTGGGCCACAGAGAAAATTGAAGCAATTTCATGATGTTATGA	1528	
RESULT	8			
LOCUS	AA449300	443 bp	mRNA	EST 04-JUN-1997
DEFINITION	zx08c06.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone			
	785866 3', mRNA sequence.			
ACCESSION	AA449300			
NID	92163149			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;			
	Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;			
	Homo.			
REFERENCE	1 (bases 1 to 443)			
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.			
TITLE	WashU-Merck EST Project 1997			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Wilson RK WashU-Merck EST Project			

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 412.

FEATURES
source

1. .443
/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTCACCAATCTGAAGTGGGAGCGGCCCTTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="785866"
/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
complement(<1.>443)
/db_xref="GDB:5983485"

mRNA

BASE COUNT 153 a 47 c 83 g 160 t
ORIGIN

Query Match 21.4%; Score 434; DB 7; Length 443;
Best Local Similarity 99.5%; Pred. No. 0.00e+00;
Matches 441; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db	1	TTTTTTTTTTTTTTAGGTAAACACAGGATGTAAAGTTTATATACAGAATATAATGTTTA	60	
Cp	2017	TTTTTTTTTTTTTTAGGTAAACACAGGATGTAAAGTTTATATACAGAATATAATGTTTA	1958	
Db	61	TCTGAAATATTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACCTAC	120	
Cp	1957	TCTGAAATATTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACCTAC	1898	
Db	121	TATGTATATTACAGGTAAGCTACAATGGGTTTAAATTCGCAAAAGTTAAGTAAGAAATGTT	180	
Cp	1897	TATGTATATTACAGGTAAGCTACAATGGGTTTAAATTCGCAAAAGTTAAGTAAGAAATGTT	1838	
Db	181	TTAAACAAGGCTTAAAGTACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTT	240	
Cp	1837	TTAAACAAGGCTTAAAGTACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTT	1778	
Db	241	GAAGAAATCATGCTATAGAAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTGAGCTG	300	
Cp	1777	GAAGAAATCATGCTATAGAAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTGAGCTG	1718	
Db	301	GAATGTGATACATGTAAACAGTTTAAAGTTCCCATTTGAAGGTATAAAATGATGAATGTTGG	360	
Cp	1717	GAATGTGATACATGTAAACAGTTTAAAGTTCCCATTTGAAGGTATAAAATGATGAATGTTG-	1659	
Db	361	TAAGACTTAGACACTGGTCTCAGTCTGGAGCTGTGAAGATGTTGAGATAACAGCCAGC	420	
Cp	1658	TAAGACTTAGACACTGGTCTCAGTCTGGAGCTGTGAAGATGTTGAGATAACAGCCAGC	1599	
Db	421	TTTATCTCAACAGGGTTTGTGAC	443	
Cp	1598	TTTATCTCAACAGGGTTTGTGAC	1576	

RESULT 9 AA922948 482 bp mRNA EST 09-JUN-1998
LOCUS oh10hl1.s1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1457445 3'
DEFINITION similar to TR:008862 008862 SECRETED FRIZZLED RELATED PROTEIN

SFRP-2. [2] TR:P97299 ;, mRNA sequence.

ACCESSION AA922948
NID 93070257
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 870 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 377.
Location/Qualifiers
1. .482
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1457445"
/clone_lib="NCI-CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
BASE COUNT 103 a 119 c 127 g 133 t
ORIGIN
Query Match 21.4%; Score 433; DB 17; Length 482;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 10 AAAATATTCATGATTTTATTAGTTTGAATATTTCTACAAGATTTCGGGTGGGCTTTTCCTT 69
|||||
Cp 1406 AAAATATTCATGATTTTATTAGTTTGAATATTTCTACAAGATTTCGGGTGGGCTTTTCCTT 1347
Db 70 TAGGTGAAACAGCTATCCACTCCTGTGGCTTATAACTCAGGAAATGCTGGGATGCAA 129
|||||
Cp 1346 TAGGTGAAACAGCTATCCACTCCTGTGGCTTATAACTCAGGAAATGCTGGGATGCAA 1287
Db 130 ACGTGAAAAGGCAGGGGAAGCTGCCAGGCTGAGACTGGAGCAGCTAGAGTGTGCTT 189
|||||
Cp 1286 ACGTGCAAAGGCAGGGGAAGCTGCCAGGCTGAGACTGGAGCAGCTAGAGTGTGCTT 1227
Db 190 GGGGAACGGAGCTGAGATCCCGGAGCAGAAATGGTCAGCCGTGCTCGGAGCAGGCCTG 249
|||||
Cp 1226 GGGGAACGGAGCTGAGATCCCGGAGCAGAAATGGTCAGCCGTGCTCGGAGCAGGCCTG 1167
Db 250 TCGGAGCCATCAGGATGCCGGAGCTAGCACTGCAGCTTTCGAATGCTGCGGAGATGCGC 309
|||||
Cp 1166 TCGGAGCCATCAGGATGCCGGAGCTAGCACTGCAGCTTTCGAATGCTGCGGAGATGCGC 1107
Db 310 TTGAACCTCTCTGCCCCCTTCTGCCACCCTTCACCGAGGTGATCACCAGCTCCCCACCC 369

|||||
Cp 1106 TTGAACCTCTCTGCCCCCTTCTGCCACCGCTTCAACGAGGTGATCACCAGCTCCCCACCC 1047
Db 370 TGTTCCTGTCCTCCATGACCATAGGGCGGCTTGATGTCGTTTCATCTCTCCTCACAGGTGCAC 429
|||||
Cp 1046 TGTTCCTGTCCTCCATGACCATAGGGCGGCTTGATGTCGTTTCATCTCTCCTCACAGGTGCAC 987
Db 430 TCCAAGCTGTCTTTG 444
|||||
Cp 986 TCCAAGCTGTCTTTG 972
RESULT 10
LOCUS AA431859 427 bp mRNA EST 22-MAY-1997
DEFINITION zw51g08.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
773630 3', mRNA sequence.
ACCESSION AA431859
NID 92115567
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 427)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -4lml3 fwd. ET from Amersham
High quality sequence stop: 214.
Location/Qualifiers
1. .427
/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCTTAATTTTTCCTTTTTCCTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="773630"
/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
complement(<1. .>427)
BASE COUNT 154 a 46 c 79 g 148 t
ORIGIN
mRNA
Query Match 21.0%; Score 425; DB 5; Length 427;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 426; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 TTTTATAGGTAACAGGATGTAAAGTTTATATACAGAATATAATGTTTATCTGAATA 60
|||||
Cp 2008 TTTTATAGGTAACAGGATGTAAAGTTTATATACAGAATATAATGTTTATCTGAATA 1949

Db 61 TTTACAGTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATAT 120
|||||
Cp 1948 TTTACAGTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATAT 1889
|||||
Db 121 TACAGGTAAGCTACAATGGGTTTAATTTGC AAAAGTTAAGTAAGAAATGTTTTAAACAAG 180
|||||
Cp 1888 TACAGGTAAGCTACAATGGGTTTAATTTGC AAAAGTTAAGTAAGAAATGTTTTAAACAAG 1829
|||||
Db 181 GCTTAAAGTACTCAAGTCAATATAAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATC 240
|||||
Cp 1828 GCTTAAAGTACTCAAGTCAATATAAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATC 1769
|||||
Db 241 ATGCTATAGAAATGGTTAAATGTGCTTCTTAATAAAATGGAAGTATTTAGCTGGAATGTGAT 300
|||||
Cp 1768 ATGCTATAGAAATGGTTAAATGTGCTTCTTAATAAAATGGAAGTATTTAGCTGGAATGTGAT 1709
|||||
Db 301 ACATGTAACAGTTTAAGTTCCTCCATTGAAGGTATAAAATGATGAATTTGTTGAAGACTTAG 360
|||||
Cp 1708 ACATGTAACAGTTTAAGTTCCTCCATTGAAGGTATAAAATGATGAATTTGTTGAAGACTTAG 1649
|||||
Db 361 ACACTGAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCTCAA 420
|||||
Cp 1648 ACACTGAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCTCAA 1589
|||||
Db 421 CAGGGTT 427
|||||
Cp 1588 CAGGGTT 1582
|||||
RESULT 11
LOCUS AA927991 518 bp mRNA EST 23-JUN-1998
DEFINITION on47d08.s1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1559823 3'
similar to TR:O08862 O08862 SECRETED FRIZZLED RELATED PROTEIN
SFRP-2. [2] TR:P97299 ;, mRNA sequence.
ACCESSION AA927991
NID g3076735
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 518)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 853 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 194.
Location/Qualifiers
1. .518
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

/db_xref="taxon:9606"
/clone="IMAGE:1559823"
/clone_lib="NCI_CGAP_Co8"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
BASE COUNT 112 a 123 c 141 g 142 t
ORIGIN
Query Match 20.7%; Score 420; DB 18; Length 518;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 425; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Db 10 AAAATATTCATGATTTTATTAGTTTGAATATTTCTACAAAGATTCGGGTGGCTTTTCCTT 69
|||||
Cp 1406 AAAATATTCATGATTTTATTAGTTTGAATATTTCTACAAAGATTCGGGTGGCTTTTCCTT 1347
|||||
Db 70 TAGGTGAAAAACAGCTATCCACTCTCTGTGGCCTTATAACTCAGGAAATGCTGGGGATGCAA 129
|||||
Cp 1346 TAGGTGAAAAACAGCTATCCACTCTCTGTGGCCTTATAACTCAGGAAATGCTGGGGATGCAA 1297
|||||
Db 130 ACCTGCAAAAGGCAGGGGGAAGCTGCCAGGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 189
|||||
Cp 1286 ACCTGCAAAAGGCAGGGGGAAGCTGCCAGGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 1227
|||||
Db 190 GGGGAACGGGAGCTGAGATCCCGGAGCAGAAATGGTCAGCGTGCTCTGGAGCAGGCCTG 249
|||||
Cp 1226 GGGGAACGGGAGCTGAGATCCCGGAGCAGAAATGGTCAGCGTGCTCTGGAGCAGGCCTG 1167
|||||
Db 250 TCGGAGCCATCAGGATCCGGGACTAGCACTGCAGCTTGGGATGCTCGGGGAGATGCGC 309
|||||
Cp 1166 TCGGAGCCATCAGGATCCGGGACTAGCACTGCAGCTTGGGATGCTCGGGGAGATGCGC 1107
|||||
Db 310 TTGAACCTCTCTCTGCCCTTCTGCCACCGCTTCGACGAGGTGATCACCAGCTCCCCACCC 369
|||||
Cp 1106 TTGAACCTCTCTCTGCCCTTCTGCCACCGCTTCACCGAGGTGATCACCAGCTCCCCACCC 1047
|||||
Db 370 TGTTTCTGTCCCATGACCAGATAGGGGCGCGTTTATGTCTGTTTCATCTACTCACAGGTGCAC 429
|||||
Cp 1046 TGTTTCTGTCCCATGACCAGATAGGGGCGCGTTGATGTCTGTTTCATCTCTCACAGGTGCAC 987
|||||
Db 430 TGAAGCTGT 439
|| |||||
Cp 986 TGAAGCTGT 977
|||||
RESULT 12
LOCUS N64004 446 bp mRNA EST 01-MAR-1996
DEFINITION YY64d09.s1 Homo sapiens cDNA clone 278321 3'.
ACCESSION N64004
NID g1211833
KEYWORDS EST.
SOURCE human clone=278321 primer=ml3 -40 forward library=Soares multiple
sclerosis 2NbHMSp vector=pT7T3D (Pharmacia) with a modified
polylinker V_TYPE: phagemid host=DH10B (ampicillin resistant)
Rsitel=Not I Rsitel2=Eco RI 46 year old male. 1st strand cDNA was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCATTTTCTTTTCTTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T vector (Pharmacia). Library went
through one round of normalization to a Cot = 5. Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4
multiple sclerosis lesions from one patient was kindly provided by
Dr. Kevin G. Becker (NINDS/NIH).
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 446)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,

Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
1. .446
/organism="Homo sapiens"
/clone="278321"
<1. .>446
157 a 51 c 87 g 149 t 2 others
BASE COUNT
ORIGIN
Query Match 20.5%; Score 416; DB 22; Length 446;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 440; Conservative 0; Mismatches 1; Indels 4; Gaps 4;
Db 2 TAGGTAAACAGGATGTAAGTTTATATACAAGAATATAATGTTTATCTGAAATATTAC 61
|||||
Cp 2003 TAGGTAAACAGGATGTAAGTTTATATACAAGAATATAATGTTTATCTGAAATATTAC 1944
Db 62 AGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGFAAACTACTATGTATATTACAG 121
|||||
Cp 1943 AGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGFAAACTACTATGTATATTACAG 1884
Db 122 GTAAGCTACAATGGTTTAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCCTA 181
|||||
Cp 1883 GTAAGCTACAATGGTTTAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCCTA 1824
Db 182 AAGTACTCAAGTCAATTATAAAATTTATATCTTTTGGCCTTTTACTTGAAGAAATCATGCT 241
|||||
Cp 1823 AAGTACTCAAGTCAATTATAAAATTTATATCTTTTGGCCTTTTACTTGAAGAAATCATGCT 1764
Db 242 ATAGAAATGGTTAATGTGCTTCTAATAAATGGAAGTATGTAGCTGGAATGTGATACATG 301
|||||
Cp 1763 ATAGAAATGGTTAATGTGCTTCTAATAAATGGAAGTATGTAGCTGGAATGTGATACATG 1704
Db 302 TAACAGTTTAAGTTCCTCCATGAAGGTATAAAATNGATGAATGTTGTAAGACCTTAGACA 361
|||||
Cp 1703 TAACAGTTTAAGTTCCTCCATGAAGGTATAAAAT-GATGAATGTTGTAAGAC-TTAGACA 1646
Db 362 CTGGAGTCTCAGTCTGGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCTCAAC 421
|||
Cp 1645 CTG-AGTCTCAGTCTGG-AGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCTCAAC 1588
Db 422 AGGGTTTGTGACCCNCAAGTTTGGG 446
|||||
Cp 1587 AGGGTTTGTGACCCACAAAGTTTGGG 1563
RESULT 13
LOCUS w77968 462 bp mRNA EST 17-OCT-1996
DEFINITION zd70ell1.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
346028 5', mRNA sequence.
ACCESSION w77968
NID g1388502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 462)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfig,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 424.
Location/Qualifiers
1. .462
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGGCATCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."
/db_xref="taxon:9606"
/clone="346028"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1. .>462
128 a 133 c 106 g 91 t 4 others
BASE COUNT
ORIGIN
mRNA
Query Match 20.4%; Score 413; DB 24; Length 462;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 449; Conservative 0; Mismatches 3; Indels 6; Gaps 6;
Db 6 TCGCTCTCGCCCCCGTCTGCCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCG 65
|||||
QY 544 TCGCTCTCGCCCCCGTCTGCCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCG 603
Db 66 CTCTGN-TGCAGGTGAAGGACCGCTGCGCCCCCGGTCATGTCGCCCTTCGN-TTCCCCTGG 123
|||||
QY 604 CTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCCGGTCATGTCGCCCTTCGGCTTCCCCTGG 663
Db 124 CCCGACATGCTTGAGTGGCGACCGTTTCCCCCAGGACAACGACCTTTCATCCCCCTCCGC 183
|||||
QY 664 CCCGACATGCTTGAGTGGCGACCGTTTCCCCCAGGACAACGACCTTTCATCCCCCTC-GC 722
Db 184 TAGCAGN-ACCACCTCTGCCAGCCACCAGGAGAGCTCCAAAGGTATGTGAAGCCTGCAA 242
|||||
QY 723 TAGCAGCACCACCTCTGCCAGCCACCAGGAGAGCTCCAAAGGTATGTGAAGCCTGCAA 782
Db 243 AAATAAAAATGATGATGACAACGACATAATGGAACGCTTTTGTAATAATGATTTTGCACT 302
|||||
QY 783 AAATAAAAATGATGATGACAACGACATAATGGAACGCTTTTGTAATAATGATTTTGCACT 842
Db 303 GAAATAAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAATATCATCTCTGGAAGA 362
|||||
QY 843 GAAATAAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAATATCATCTCTGGA-GA 901

Db 363 CCAAGACGAAGACCATTACAAAGCTGAACGGTGTGTCCGAAAGGGACCTTGAAGAAATCG 422
|||||
QY 902 CCAAGACGAAGACCATTACAAAGCTGAACGGTGTGTCCGAAAGGGACCT-GAAGAAATCG 960
Db 423 GTGCTGTGGCTCAAAGACAGACTTGCAGTGCACCTGTGA 460
|||||
QY 961 GTGCTGTGGCTCAAAGACAGACTTGCAGTGCACCTGTGA 998

RESULT 14
LOCUS AA449749 409 bp mRNA EST 04-JUN-1997
DEFINITION zx07e10.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
785802 3', mRNA sequence.
ACCESSION AA449749
NID g2163499
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 409)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 408.
Location/Qualifiers
1. .409
/organism="Homo sapiens"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - Oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="785802"
/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
complement(<1. .>409)
/db_xref="GDB:5983421"

mRNA
BASE COUNT 148 a 42 c 75 g 144 t
ORIGIN

Query Match 20.2%; Score 409; DB 7; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTTTCTAGGTAACAGGATGTAAAGTTTATATACAAGAATATATCTTATCTGAAA 60
|||||
Cp 2010 TTTTCTAGGTAACAGGATGTAAAGTTTATATACAAGAATATATCTTATCTGAAA 1951
|||||
Db 61 TATTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACACTATGTAT 120
|||||

Cp 1950 TATTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACACTATGTAT 1891
Db 121 ATTACAGGTAAGCTACAATGGGTTTAATTTGCAAAAAGTTAAGTAAGAAATGTTTAAACA 180
|||||
Cp 1890 ATTACAGGTAAGCTACAATGGGTTTAATTTGCAAAAAGTTAAGTAAGAAATGTTTAAACA 1831
Db 181 AGGCTTAAAGTACTCAAGTCAATTATATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAA 240
|||||
Cp 1830 AGGCTTAAAGTACTCAAGTCAATTATATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAA 1771
Db 241 TCATGCTATAGAAAATGGTTAATGTGCTCTTAATAAATGGAAGTATTGTAGCTGGAATGTG 300
|||||
Cp 1770 TCATGCTATAGAAAATGGTTAATGTGCTCTTAATAAATGGAAGTATTGTAGCTGGAATGTG 1711
Db 301 ATACATGTACAGTTTAAAGTTCCCATTTGAGCTATAGATGATGAATTTGTTGAAGACTT 360
|||||
Cp 1710 ATACATGTACAGTTTAAAGTTCCCATTTGAGCTATAGATGATGAATTTGTTGAAGACTT 1651
Db 361 AGACACTGAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATAACAGCC 409
|||||
Cp 1650 AGACACTGAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATAACAGCC 1602

RESULT 15
LOCUS N32415 445 bp mRNA EST 10-JAN-1996
DEFINITION YW83c06.s1 Homo sapiens cDNA clone 258826 3'.
ACCESSION N32415
NID g1152814
KEYWORDS EST.
SOURCE human clone=258826 primer=m13 -40 forward library=Soares placenta
8to9weeks 2NbHP8to9W vector=PT7T3D (Pharmacia) with a modified
polylinker host=DH10B (ampicillin resistant) Rsite1=Not I
Rsite2=Eco RI two placentae: One from 8 weeks and another from 9
weeks post conception. 1st strand cDNA was primed with a Not I -
oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCGGATTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). Library
constructed by Bento Soares and M. Fatima Bonaldo.
Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 445)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 314
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
1. .445
/organism="Homo sapiens"
/clone="258826"
<1. .>445

FEATURES
Source
mRNA
BASE COUNT 156 a 55 c 87 g 147 t
ORIGIN

Query Match 20.1%; Score 408; DB 21; Length 445;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 434; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 AAAACAGGATGTAAAGTTTATATACAAGATATAATGTTTATCTGAAATATTACAGTGT 60
|||||
Cp 1998 AAAACAGGATGTAAAGTTTATATACAAGATATAATGTTTATCTGAAATATTACAGTGT 1939
|||||

Db 61 TGGTTAAAGCAATATTTTACAACCTTTTAAAGTAAACTACTATGTATATTACAGGTAAG 120
|||||
Cp 1938 TGGTTAAAGCAATATTTTACAACCTTTTAAAGTAAACTACTATGTATATTACAGGTAAG 1879
|||||

Db 121 CTACAATGGGTTTAATTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAAGTA 180
|||||
Cp 1878 CTACAATGGGTTTAATTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAAGTA 1819
|||||

Db 181 CTCAGTCAATTATATAAATTTATATCTTTTGGCCTTTTACTTTGAAGAAATCATGCTATAGA 240
|||||
Cp 1818 CTCAGTCAATTATATAAATTTATATCTTTTGGCCTTTTACTTTGAAGAAATCATGCTATAGA 1759
|||||

Db 241 AATGGTTAATGCTTCTTAATAAATGGAAGTATGTAGCTGGAATGTGATACATGTAACA 300
|||||
Cp 1758 AATGGTTAATGCTTCTTAATAAATGGAAGTATGTAGCTGGAATGTGATACATGTAACA 1699
|||||

Db 301 GTTTAAGTCCCATTTGAAGGTATATAAATGATGAATTTGTTGAAGACTTAGACACTGGAGT 360
|||||
Cp 1698 GTTTAAGTCCCATTTGAAGGTATATAAATGATGAATTTGTTGAAGACTTAGACACTG-AGT 1640
|||||

Db 361 CTCAGTCTGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCCCCAACAGGGGT 420
|||||
Cp 1639 CTCAGTCTGG-AGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCTC-AACAGGG-T 1583
|||||

Db 421 TTGTGACCCACAAGTTTGG 440
|||||
Cp 1582 TTGTGACCCACAAGTTTGG 1563
|||||

RESULT 16
LOCUS AA449032 562 bp mRNA EST 04-JUN-1997
DEFINITION zx07e10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 785802 5' similar to TR:G310115 G310115 FRIZZLED PROTEIN HOMOLOG 2
PRECURSOR ; mRNA sequence.

ACCESSION AA449032
NID g2163052
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 562)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 467.
Location/Qualifiers
1. .562
/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

FEATURES
Source

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGCCCTTAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo.
/db_xref="taxon:9606"
/clone="785802"
/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
<1. .>562
/db_xref="GDB:5983421" 98 t

BASE COUNT 75 a 232 c 157 g 98 t
ORIGIN

Query Match 20.1%; Score 407; DB 7; Length 562;
Best Local Similarity 98.2%; Pred. No. 0.00e+00;
Matches 440; Conservative 0; Mismatches 3; Indels 5; Gaps 5;

Db 1 GCTCATTTCTGCTCCCCGGGTGGAGCCCCCGGAGCTGCGCGGGGCTTGCAGCGC-TC 59
|||||
QY 22 GCTCATTTCTGCTCCCCGGGTGGAGCCCCCGGAGCTGCGCGGGGCTTGCAGCGCGCTC 81
|||||

Db 60 GCCCGCGTGTCTCCCGGTGTCCCGCTTCTCCGCGCCCGCCAGCCCGCGG-TGCCAGCTTTT 118
|||||
QY 82 GCCCGCGC-TCCTCCCGGTGTCCCGCTTCTCCGCGCCCGCCAGCCCGCGGCTGCCAGCTTTT 140
|||||

Db 119 CGGGGCCCCGAGTCGCACCCAGCGAAGAGAGCGGGCCCCGGGACAAAGCTCGAACTCCGGCC 178
|||||
QY 141 CGGGGCCCCGAGTCGCACCCAGCGAAGAGAGCGGGCCCCGGGACAAAGCTCGAACTCCGGCC 200
|||||

Db 179 GCCTCGCCCTTCCCGGCTCCGCTCCCTCTGCCCCCTCGGG-TCGCGCGCCCGACGATGCT 237
|||||
QY 201 GCCTCGCCCTTCCCGGCTCCGCTCCCTCTGCCCCCTCGGGTTCGCGCGCCCGACGATGCT 260
|||||

Db 238 GCAGGGCCCTGGCTCGCTGCTGCTCTTCTCCTCGCTCGCACTGCTGCCTGGGCTCGGC 297
|||||
QY 261 GCAGGGCCCTGGCTCGCTGCTGCTCTTCTCCTCGCTCGCACTGCTGCCTGGGCTCGGC 320
|||||

Db 298 GCGCGGGCTCTTCTCTTTGGCCAGCCCCGACTTCTCTACAAGCGCAG-AATTGCAAGCC 356
|||||
QY 321 GCGCGGGCTCTTCTCTTTGGCCAGCCCCGACTTCTCTACAAGCGCAGCAATTGCAAGCC 380
|||||

Db 357 CATCCCTGCCAACCTGCAGCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAA 416
|||||
QY 381 CATCCCGGCCAACCTGCAGCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAA 440
|||||

Db 417 CCTGCTGGCCACGAGACCATGAAGCAG 444
|||||
QY 441 CCTGCTGGCCACGAGACCATGAAGGAG 468
|||||

RESULT 17
LOCUS AA505877 405 bp mRNA EST 20-AUG-1997
DEFINITION ni01a05.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:966704, mRNA sequence.

ACCESSION AA505877
NID 92242014
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 405)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image.html

Insert Length: 935 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 315.
Location/Qualifiers
1. .405
/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dt) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. This library is the normalized version of NCI-CGAP_Brl.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:966704"
/clone_lib="NCI-CGAP_Br2"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
<1. .>405

149 a 43 c 74 g 139 t

mRNA

BASE COUNT

ORIGIN

Query Match 20.0%; Score 405; DB 8; Length 405;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 AAACACGAGTGTAAGTTTATATACAAGAAATATAATGTTTATCTGAAATATTTACAGTGT 60
|||||

Cp 1998 AAACACGAGTGTAAGTTTATATACAAGAAATATAATGTTTATCTGAAATATTTACAGTGT 1939
|||||

Db 61 TGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGGTAAG 120
|||||

Cp 1938 TGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGGTAAG 1879
|||||

Db 121 CTACAATGGGTTTAATTTGCAAAAAGTTAAGTAAGAAATGTTTTAAACAAGGCTTAAAGTA 180
|||||

Cp 1878 CTACAATGGGTTTAATTTGCAAAAAGTTAAGTAAGAAATGTTTTAAACAAGGCTTAAAGTA 1819
|||||

Db 181 CTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGA 240
|||||

Cp 1818 CTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGA 1759
|||||

Db 241 AATGGTTAATGTGCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAACA 300
|||||

Cp 1758 AATGGTTAATGTGCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAACA 1699
|||||

Db 301 GTTTAAGTTCCCATTTGAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACTGAGTC 360
|||||

Cp 1698 GTTTAAGTTCCCATTTGAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACTGAGTC 1639
|||||

Db 361 TCAGTCTGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTAT 405
|||||

Cp 1638 TCAGTCTGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTAT 1594
|||||

RESULT 18 AA976403 408 bp mRNA EST 22-MAY-1998
LOCUS

DEFINITION

oq68f09.s1 NCI-CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1591529 3', mRNA sequence.

ACCESSION

AA976403

NID

g3152195

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 408)

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Seg primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 402.
Location/Qualifiers
1. .408
/organism="Homo sapiens"
/note="Organ: kidney; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled kidney tumors. 5' adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.0 kb."
/db_xref="taxon:9606"
/clone="IMAGE:1591529"
/clone_lib="NCI-CGAP_Kid6"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="SOLR (kanamycin resistant)"

95 a 108 c 90 g 115 t

BASE COUNT

ORIGIN

Query Match 19.8%; Score 402; DB 17; Length 408;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 408; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 GAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGTAGTCCCGGCATCCTG 60
|||||

QY 1097 GAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGTAGTCCCGGCATCCTG 1156
|||||

Db 61 ATGCTCCGACAG-CCTGCTCCAGACGCGCTGACCATTTCTGCTCCGGGATCTCAGCT 119
|||||

QY 1157 ATGGTCCGACAGGCGCTGCTCCAGACGCGCTGACCATTTCTGCTCCGGGATCTCAGCT 1216
|||||

Db 120 CCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGCAGCTTCCCCCTGCC 179
|||||

QY 1217 CCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCGAGCTTCCCCCTGCC 1276
|||||

Db 180 TTTTGCACGTTTGCATCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTGGATAGCTG 239
|||||

QY 1277 TTTTGCACGTTTGCATCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTGGATAGCTG 1336
|||||

Db 240 TTTTCACCTAAAGGAAAGCCCAACCCGAATCTTGTAGAAATATATCAAACTAATAAAATCA 299
|||||

QY 1337 TTTTCACCTAAAGGAAAGCCCAACCCGAATCTTGTAGAAATATATCAAACTAATAAAATCA 1396
|||||

Db 300 TGAATATTTTATGAAGTTTAAAAATAGCTCACCTTTAAAGCTAGTTTGAATAGGTGCAA 359
|||||

QY 1397 TGAATATTTTATGAAGTTTAAAAATAGCTCACCTTTAAAGCTAGTTTGAATAGGTGCAA 1456
|||||

Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 372; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 51 CTAGTCCCGGATCCTGATGGCTCCGACAG-CCTGCTCCAGAGCACGGCTGACCATTTCT 109
QY 1140 CTAGTCCCGGATCCTGATGGCTCCGACAGGCTGCTCCAGAGCACGGCTGACCATTTCT 1199
Db 110 GCTCCGGGATCTCAGCTCCCGTTTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTG 169
QY 1200 GCTCCGGGATCTCAGCTCCCGTTTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTG 1259
Db 170 GGCAGCTTCCCTGCTTTTGCACGTTTGCATCCCTCCAGCATTTCTCGAGTTATAAGGCC 229
QY 1260 GGCAGCTTCCCTGCTTTTGCACGTTTGCATCCCTCCAGCATTTCTCGAGTTATAAGGCC 1319
Db 230 ACAGGAGTGGATAGCTGTTTTCACCTAAAGGAAAGCCACCCGAATCTTGAGAAATAT 289
QY 1320 ACAGGAGTGGATAGCTGTTTTCACCTAAAGGAAAGCCACCCGAATCTTGAGAAATAT 1379
Db 290 TCAAACTAATAAATCATGAATATTTTATGAGTTTAAAAATAGCTCAGCTTAAAGCTA 349
QY 1380 TCAAACTAATAAATCATGAATATTTTATGAGTTTAAAAATAGCTCAGCTTAAAGCTA 1439
Db 350 GTTTTGAATAGGTGCAACTGTGACTTGGGTCTGGTGGTTGTTGTTTGTGTTGAGTC 409
QY 1440 GTTTTGAATAGGTGCAACTGTGACTTGGGTCTGGTGGTTGTTGTTGTTGTTGAGTC 1499
Db 410 AGCTGATTTTCAC 422
QY 1500 AGCTGATTTTCAC 1512
RESULT 21
LOCUS AA105749 458 bp mRNA EST 04-FEB-1997
DEFINITION ml84a03.r1 Stratagene mouse kidney (#937315) Mus musculus cDNA clone 518668 5' similar to TR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR. ;, mRNA sequence.
ACCESSION AA105749
NID g1654838
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 458)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:312516
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 457.
Location/Qualifiers
1. .458
/organism="Mus musculus"
/strain="C57/Bl6"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCCGGCACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3"
/db_xref="taxon:10090"
/clone="518668"
/clone_lib="Stratagene mouse kidney (#937315)"
/sex="females"
/dev_stage="4 weeks"
/lab_host="SOLR (kanamycin resistant)"
<1. .>458
mRNA
BASE COUNT 119 a 140 c 116 g 83 t
ORIGIN
Query Match 18.1%; Score 366; DB 25; Length 458;
Best Local Similarity 91.8%; Pred. No. 0.00e+00;
Matches 424; Conservative 0; Mismatches 34; Indels 4; Gaps 4;
Db 1 AGCAGGGCGCGCATGGATTCCGCTGGTTCATGAAGCAGTCCACCCGGACACCAAGAAGT 60
QY 476 AGCAGGGCGCGCTTGGATCCCGCTGGTTCATGAAGCAGTCCACCCGGACACCAAGAAGT 535
Db 61 TCCTGTGCTCGCTCTTCGCCCCCTGTCTGTCTCGACGACCTAGATGAGACCATCCAGCGT 120
QY 536 TCCTGTGCTCGCTCTTCGCCCCCTGTCTGTCTCGATGACCTAGACGACCATCCAGCAT 595
Db 121 GTCACTCGCTCTG-GTGCAGGTGAAGACCGCTGCGCCCCGGTTCATGTCCGCCCTTCG-CT 178
QY 596 GCCACTCGCTCTGCGTGCAGGTGAAGACCGCTGCGCCCCGGTTCATGTCCGCCCTTCGCT 655
Db 179 TCCCTGGCCAGACATGCTGGAGTGCAGCCGTTTCCCGCAG-ACAACGACCTCTGCATCC 237
QY 656 TCCCTGGCCGACATGTTGAGTGCAGCCGTTTCCCGCAGCAACGACCTTTGCATCC 715
Db 238 CCCTCGCTAGTAG-GACCACCTCCTGCCGGCCACAGAGGAAGCTCCCAAGGTGTGTGAAG 296
QY 716 CCCTCGTAGCAGCGACCACTCCTCTGCCAGCCACCAGGAAAGCTCCAAAGGTATGTGAAG 775
Db 297 CCTGCAAAACCAAGAATGAGGACGACACACATCATCGAAACCCCTTTGTAAAAATGACT 356
QY 776 CCTGCAAAATAAAAAATGATGATGACAACGACATAATGGAACGCTTTGTAAAAATGATT 835
Db 357 TCGCACTGAAAAATCAAAGTGAAGGAGATAACGTACATCAACAGAGACACCAAGATCATCC 416
QY 836 TTGCACTGAAAAATAAAGTGAAGGAGATAAACCTACATCAACCGAGATACCAAAATCATCC 895
Db 417 TGGAGACAAAGAGCAAGACCATTTACAAGCTGAACGGCGTGT 458
QY 896 TGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGT 937
RESULT 22
LOCUS AA429960 378 bp mRNA EST 16-OCT-1997
DEFINITION zw58g10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 774306 5', mRNA sequence.
ACCESSION AA429960
NID g2113134
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 378)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 352.

FEATURES
source
1. .378

/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCTTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="774306"

/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"

BASE COUNT 115 a 101 c 88 g 74 t

Query Match 17.9%; Score 362; DB 10; Length 378;
Best Local Similarity 99.2%; Pred. NO. 0.00e+00;
Matches 370; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 7 CCCGGTCATGTCGCCCTTCG-CTTCCCCTGGCCCGACATGCTTGAGTGGACCGTTTCCC 65
|||||

QY 633 CCCGGTCATGTCGCCCTTCGGCTTCCCCTGGCCCGACATGCTTGAGTGGACCGTTTCCC 692

Db 66 CCAGGACAACGACCTTTGCATCCCCCTCGCPAGCAGCGACCCTCCTGCCAGCCACCGA 125
|||||

QY 693 CCAGGACAACGACCTTTGCATCCCCCTCGCPAGCAGCGACCCTCCTGCCAGCCACCGA 752

Db 126 GGAAGCTCCAAAGGTATGTGAAGCCTGC AAAAATAAAAAATGATGATGACACGACATAAT 185
|||||

QY 753 GGAAGCTCCAAAGGTATGTGAAGCCTGC AAAAATAAAAAATGATGATGACACGACATAAT 812

Db 186 GGAAACGCTTTGTAAAAATGATTTGCACCTGAAAATAAAAGTGAAGGAGATAACCTACAT 245
|||||

QY 813 GGAAACGCTTTGTAAAAATGATTTGCACCTGAAAATAAAAGTGAAGGAGATAACCTACAT 872

Db 246 CAACCGAGATACAAAATCATCCTGGAGACCAAGACCAACCATTTACAGCTGAACGG 305
|||||

QY 873 CAACCGAGATACAAAATCATCCTGGAGACCAAGACCAACCATTTACAGCTGAACGG 932

Db 306 TGTGTCGGAATAGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCAC 365
|||||

QY 933 TGTGTCGGAAGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCAC 992

Db 366 CTGTGAGGAGATG 378

QY 993 CTGTGAGGAGATG 1005

RESULT 23
LOCUS AA024772 433 bp mRNA EST 13-AUG-1996
DEFINITION ze76g10.s1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
364962 3', mRNA sequence.

ACCESSION AA024772

NID g1489678

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 433)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,

TITLE
JOURNAL
COMMENT

Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 289.

FEATURES
source
1. .433

/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCATCTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."
/db_xref="taxon:9606"
/clone="364962"
/clone_lib="Soares fetal heart NBHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
complement(<1.>433)

mRNA 150 a 51 c 87 g 141 t 4 others
BASE COUNT
ORIGIN

Query Match 17.9%; Score 363; DB 23; Length 433;
Best Local Similarity 99.2%; Pred. No. 0.00e+00;
Matches 369; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 GTAAAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAATATTACAGT 60
|||||

Cp 2000 GTAAAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAATATTACAGT 1941

Db 61 GTTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACACTACTATGTATATTACAGGTA 120
|||||

Cp 1940 GTTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACACTACTATGTATATTACAGGTA 1881

Db 121 AGCTACAATGGGTTTAAATTGCAAAAAGTTAAGTAAGAAATGTTTTTAAACAAGGCTTAAAG 180
|||||

Cp 1880 AGCTACAATGGGTTTAAATTGCAAAAAGTTAAGTAAGAAATGTTTTTAAACAAGGCTTAAAG 1821

Db 181 TACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATA 240
|||||

Cp 1820 TACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATA 1761

Db 241 GAAATGGTTAATGTGCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAA 300
|||||

Cp 1760 GAAATGGTTAATGTGCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAA 1701

Db 301 CAGTTTAAGTTCCTCATTAAGGTTATAAAATGATGAATGTTGTTAAGACTTAGACACNGGA 360
|||||

Cp 1700 CAGTTTAAGTTCCTCATTAAGGTTATAAAATGATGAATGTTGTTAAGACTTAGACACTG-A 1642

Db 361 GTCTCAGTCNGG 372

Cp 1641 GTCTCAGTCTGG 1630

RESULT 24 H87071 426 bp mRNA EST 21-NOV-1995
LOCUS ys74d07.r1 Homo sapiens cDNA clone 220525 5' similar to SP:A45054
DEFINITION A45054 F2-1-PUTATIVE INTERCELLULAR SIGNAL TRANSDUCER OR TRANSMITTER
- ;
ACCESSION H87071
NID 91068650
KEYWORDS EST.
SOURCE human clone=220525 primer=M13RP1 library=Soares retina N2b4HR
vector=pt7T3D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA
was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pt7T3 vector (Pharmacia). The retinas were
obtained from a 55 year old Caucasian male and total cellular
poly(A)+ RNA was extracted 6 hrs after their removal. The retina
RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento Soares and
M.Fatima Bonaldo.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 355
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES
source Location/Qualifiers
1. .426
/organism="Homo sapiens"
/clone="220525"
<1. .>426
BASE COUNT 96 a 128 c 110 g 83 t 9 others
ORIGIN
Query Match 17.7%; Score 359; DB 21; Length 426;
Best Local Similarity 94.8%; Pred. No. 0.00e+00;
Matches 399; Conservative 0; Mismatches 17; Indels 5; Gaps 5;
Db 2 ACGGCATNGAATACCAGACATCGCGGTGCCCAACCTGCTGGGCCACGAGACCATGAAGG 61
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 407 ACGGCATCGAATACCAGACATCGCGGTGCCCAACCTGCTGGGCCACGAGACCATGAAGG 466
Db 62 AGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCGGACA 121
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 467 AGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCGGACA 526
Db 122 CCAAGAAAGTTCTGTGCTCGCTCTTCGCCCGCCCGTCTGCCTCGATGACCTAGACGAGACCA 181
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 527 CCAAGAAAGTTCTGTGCTCGCTCTTCGCCCGCCCGTCTGCCTCGATGACCTAGACGAGACCA 586
Db 182 TCCAGCCATGCCACTCGCTCTG -GTGCAGGTGAAAGGACCGCTGCGCCCCGGTCAATGTCC 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 587 TCCAGCCATGCCACTCGCTCTGCTGTCAGGTGAA -GGACCGTGGCCCCCGGTCAATGTCC 645
Db 241 GCCTTCGGTTTCCCTGGCCCGACATGCTTGAGTGCAGCCGTTTCCCGCCAGGACAACGAC 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 646 GCCTTCGGTTTCCCTGGCCCGACATGCTTGAGTGCAGCCGTTTCCCGCCAGGACAACGAC 705
Db 301 CTTTGCATCCCCCTNGNTAGCAGGANCCACCTTCTTGTGTCAGCCACCGAGGAAGTTNCAA 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 706 CTTTGCATCCCCCTCGTAGCAGGACCACCTCCT -GCCAGCCACCGAGGAAGTCCCAA 764
Db 361 GGTATGTGAAGCTGNAAAAATTAAAAATGATGATGAACAACGACATATATGGGAAACGNTT 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 765 GGTATGTGAAGCTGTCAAAATAATAAATGATGATGA -CAACGACATAATGG-AAACGCTT 822
Db 421 T 421
QY 823 T 823
RESULT 25
LOCUS AA295018 359 bp mRNA EST 18-APR-1997
DEFINITION EST100474 Pancreas tumor I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA295018
NID g1947649
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other_ESTs: THC180057
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1. .359
/organism="Homo sapiens"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
/db_xref="ATCC (inhost):190769"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"


```

mRNA
BASE COUNT      106 a      96 c      82 g      68 t      7 others
ORIGIN
/dev_stage="adult"
<1. .>359
Query Match      17.3%; Score 350; DB 4; Length 359;
Best Local Similarity 97.8%; Pred. No. 0.00e+00;
Matches 351; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 1 CGTGCAGTGAAGGACCGCTGGCCCCCGGTCATGTCGGCNTTCGGNTTCCCTGGCCCCGA 60
    |||||||
QY 609 CGTGCAGTGAAGGACCGCTGGCCCCCGGTCATGTCGGCNTTCGGNTTCCCTGGCCCCGA 668
    |||||||

Db 61 CATGCTTGANTGCGACCNCTTTCCCCAGGACAAACGACCTTTTGCAATCCCTCGCTAGCAG 120
    |||||||
QY 669 CATGCTTGAGTGCAGACCGTTTCCCCAGGACAAACGACCTTTTGCAATCCCTCGCTAGCAG 728
    |||||||

Db 121 CGACCACTCCTGCCAGCCACCGAGGAGCTCCAAAGGTATGTNAAGCCTGCAAAATAA 180
    |||||||
QY 729 CGACCACTCCTGCCAGCCACCGAGGAGCTCCAAAGGTATGTGAAGCCTGCAAAATAA 788
    |||||||

Db 181 AANTGATGATGACAACGACATAATGGAACGCTTTGTAAAAATNATTTGCACTGAAAAT 240
    |||||||
QY 789 AAATGATGATGACAACGACATAATGGAACGCTTTGTAAAAATGATTTGCACACTGAAAAT 848
    |||||||

Db 241 AAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATCTGGAGACCAAGAG 300
    |||||||
QY 849 AAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATCTGGAGACCAAGAG 908
    |||||||

Db 301 CAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGGAATCGGTCTGT 359
    |||||||
QY 909 CAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGGAATCGGTCTGT 967
    |||||||

RESULT 26
LOCUS      AI047549      449 bp      mRNA      EST      08-JUL-1998
DEFINITION      uh80b08.r1 Soares mouse urogenital ridge NMUR Mus musculus cDNA
                  clone 1763991 5', similar to TR:O08862 O08862 SECRETED FRIZZLED
                  RELATED PROTEIN SFRP-2. [2] TR:P97299 ;, mRNA sequence.
ACCESSION      AI047549
NID            93295836
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 449)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
                  Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
                  Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
                  Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
                  Waterston,R.
TITLE      The WashU-HMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
                  WashU-HMI Mouse EST Project
                  Washington University School of Medicinep
                  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
                  Tel: 314 286 1800
                  Fax: 314 286 1810
                  Email: mouseest@watson.wustl.edu
                  This clone is available royalty-free through LLNL ; contact the
                  IMAGE Consortium (info@image.llnl.gov) for further information.
                  MGI:963515
                  Seq primer: -28ml3 rev2 ET from Amersham
                  High quality sequence stop: 442.
                  Location/Qualifiers
FEATURES      1. .449
                  /organism="Mus musculus"
                  /note="Organ: gonad; Vector: pT7T3D-Pac (Pharmacia) with a
                  modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                  strand cDNA was primed with a Not I - oligo(dT) primer [5'
```

```

TGTACCAATCTGAAGTGGGAGCGCGCATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library went through two rounds of normalization, and was
constructed by Bento Soares and M.Fatima Bonaldo."
/db_xref="taxon:10090"
/clone="1763991"
/clone_lib="Soares mouse urogenital ridge NMUR"
/sex="equal ratio of male:female"
/tissue_type="urogenital ridge (embryonic)"
/dev_stage="fetal, mixture of 11.5 and 12.5 dpc"
/lab_host="DH10B"

BASE COUNT      66 a      170 c      127 g      86 t
ORIGIN
Query Match      16.7%; Score 338; DB 18; Length 449;
Best Local Similarity 91.4%; Pred. No. 0.00e+00;
Matches 412; Conservative 0; Mismatches 32; Indels 7; Gaps 7;

Db 1 TCCGGCCGCTCTCGCTTCCCGCTCGGCTCCCTCGCGCCCCC-CGGGTACGCTAGTCCA 59
    |||||||
QY 194 TCCGGCCGCTCGCGCTTCCCGCTCGGCTCCCTCGCGCCCCC-CGGGTACGCGGCCA 253
    |||||||

Db 60 CGATGCCGC-GGGCCCTG-CTCGCTGCTGCTGTAGTCTCGCTCGCTCGCACTGCTGCCTGG 117
    |||||||
QY 254 CGATGCTGCAGGGCCCTGGCTGCTGCTGCTCTCTCTCGCTCGCACTGCTGCCTGG 313
    |||||||

Db 118 GCTCGGCGGTGGGATCTTCTTTCGGCCAGCCCGACTTCTCTACAAAGCGCACGAAT 177
    |||||||
QY 314 GCTCGGCGCGGGCTCTCTCTTTGGCCAGCCCGACTTCTCTACAAAGCGCACGAAT 373
    |||||||

Db 178 GCAAGCCCATCCCGCCCAACCTGCAGCTGTGC-ACGGCATCGAGTACCAGAACATGCGGC 236
    |||||||
QY 374 GCAAGCCCATCCCGCCCAACCTGCAGCTGTGCCACGCGATCGAATACCAGAACATGCGGC 433
    |||||||

Db 237 TGCCCAACCTGCTGGGCCACGAGACCATGAAGGAGTGCTGGAGCAGGGCGGCTGGAG 296
    |||||||
QY 434 TGCCCAACCTGCTGGGCCACGAGACCATGAAGGAGTGCTGGAGCAGGGCGGCTTGG 493
    |||||||

Db 297 TTCCGCTGGTCATGAAGCAGTGCCACCCGGACACCAAGAAGTTCCTGTGCTCGCTCTCG 356
    |||||||
QY 494 TCCCGCTGGTCATGAAGCAGTGCCACCCGGACACCAAGAAGTTCCTGTGCTCGCTCTCG 553
    |||||||

Db 357 CCCCTGTGCTCGACGACCTAGATGAGACCATCCAGCCGCTGTCACTCGCTCTG-GTGC 415
    |||||||
QY 554 CCCCGTCTGCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCGCTCTCGCTGC 613
    |||||||

Db 416 AGGTGAAGGAC-GCTGCGCCCC-GGTCATGTC 444
    |||||||
QY 614 AGGTGAAGGACCGCTGCGCCCCCGGTCATGTC 644
    |||||||

RESULT 27
LOCUS      H16121      406 bp      mRNA      EST      27-JUN-1995
DEFINITION      YL28c05.r1 Homo sapiens cDNA clone 159560 5'.
ACCESSION      H16121
NID            9880941
KEYWORDS      EST.
SOURCE      human clone=159560 library=Soares breast 3NbHst vector=pT7T3D
                  (Pharmacia) with a modified polylinker host=DH10B (ampicillin
                  resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Adult human.
                  1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
                  TGTACCAATCTGAAGTGGGAGCGCGCGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
                  double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
                  digested with Not I and cloned into the Not I and Eco RI sites of a
                  modified pT7T3 vector (Pharmacia). Library went through one round
                  of normalization to a Cot = 20. Library constructed by Bento Soares
                  and M.Fatima Bonaldo.
ORGANISM      Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
```

REFERENCE 1 (bases 1 to 406)
AUTHORS Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 321
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source Location/Qualifiers
1. .406
/organism="Homo sapiens"
/clone="159560"

BASE COUNT 130 a 94 c 103 g 77 t 2 others
ORIGIN

Query Match 16.6%; Score 337; DB 20; Length 406;
Best Local Similarity 98.1%; Pred. No. 0.00e+00;
Matches 368; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Db 1 CCCAGGACAACGACCTTTGCATCCCCCTCGCTAGCAGCANCACCTCCTGCCAGCCACCG 60
|||||
QY 692 CCCAGGACAACGACCTTTGCATCCCCCTCGCTAGCAGCAGCCACCTCCTGCCAGCCACCG 751
|||||
Db 61 AGGAAGCTCCAAAGGTATGTGAAGCCTGCAAAATATAAATGATGATGACAACGACATAA 120
|||||
QY 752 AGGAAGCTCCAAAGGTATGTGAAGCCTGCAAAATATAAATGATGATGACAACGACATAA 811
|||||
Db 121 TGGAAACGCTTTGTAAAAATGATTTTGCACTGAAATATAAAGTGAAGGAGATAACCTACA 180
|||||
QY 812 TGGAAACGCTTTGTAAAAATGATTTTGCACTGAAATATAAAGTGAAGGAGATAACCTACA 871
|||||
Db 181 TCAACCGAGATACCAAATCATCTCTGGAGACCAAGAGAACCACTTTACAAGCTGAACG 240
|||||
QY 872 TCAACCGAGATACCAAATCATCTCTGGAGACCAAGAGAACCACTTTACAAGCTGAACG 931
|||||
Db 241 GTGTGTCGGAAGGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGA 300
|||||
QY 932 GTGTGTCGGAAGGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGA 991
|||||
Db 301 CCTGTGAGGAGATGACGACATTCACACGGCCCTATCTGGTTTCATGGGGACAGGAAACAG 360
|||||
QY 992 CCTGTGAGGAGATGACGACAT-CAACGGCCCTATCTGGT-CATGGG-ACAG-AAACAG 1047
|||||
Db 361 GGTGGGGAGCTGGT 375
|||
QY 1048 GGT-GGGGAGCTGGT 1061
|||

RESULT 28
LOCUS HUM296B07B 340 bp mRNA EST 21-MAY-1996
DEFINITION Human aorta cDNA 5'-end GEN-296B07, mRNA sequence.
ACCESSION D62525
NID 9966299
KEYWORDS EST; EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).
SOURCE Homo sapiens cDNA to mRNA, clone_lib:Clontech human aorta polyA+ mRNA (#6572).
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;

REFERENCE 1 (bases 1 to 340)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,
Okuno,S., Ozaki,K., Shimizu,F., Shinada,Y., Shinomiya,H.,
Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y.,
Maekawa,H., Shin,S. and Nakamura,Y.
TITLE Unpublished(303)
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 340)
AUTHORS Fujiwara,T.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu
Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical
Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,
Japan (Tel:0886-65-2888, Fax:0886-37-1035)
COMMENT Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co.,Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.

FEATURES
source Location/Qualifiers
1. .340
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"

BASE COUNT 113 a 64 c 38 g 122 t 3 others
ORIGIN

Query Match 16.3%; Score 331; DB 6; Length 340;
Best Local Similarity 98.5%; Pred. No. 0.00e+00;
Matches 335; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

Db 1 CATCTTCATCAGCTCCAGACTGAGACTCAGTGTCTAAGTCTTACAACAATTCATCATTTT 60
|||||
QY 1616 CATCTTCATCAGCTCCAGACTGAGACTCAGTGTCTAAGTCTTACAACAATTCATCATTTT 1675
|||||
Db 61 ATACCTTCAATGGGAACCTTAAACTGTTACATGTATCATCTCCAGCTACAATACTTCCAT 120
|||||
QY 1676 ATACCTTCAATGGGAACCTTAAACTGTTACATGTATCATCTCCAGCTACAATACTTCCAT 1735
|||||
Db 121 TTATTAGAGGCACATTAAACCATTCTATAGCATGATTTCTTCAAGTAAAGGCAAAAGAT 180
|||||
QY 1736 TTATTAGAAGCACATTAAACCATTCTATAGCATGATTTCTTCAAGTAAAGGCAAAAGAT 1795
|||||
Db 181 ATAAATTTTATAATTGACTTGAGTACTTTTAAGCCCTTGTTTAAACAATTTCTTACTTAACT 240
|||||
QY 1796 ATAAATTTTATAATTGACTTGAGTACTTTAAGCCCTTGTTTAAACAATTTCTTACTTAACT 1855
|||||
Db 241 TTTGGCAAATTAAACCCATTGTAGCTTACCTGTAATATACATAGTRGTTTACCTTTAAAA 300
|||||
QY 1856 TTTG-CAAAATTAACCCATTGTAGCTTACCTGTAATATACATAGTAGTTTACCTTTAAAA 1914
|||||
Db 301 GTTGTAATAARTATTGCTTTRACCAACACTGTAAATATTTC 340
|||||
QY 1915 GTTGTAATAATATTGCTTTTAAACCAACACTGTAAATATTTC 1954
|||||

RESULT 29
LOCUS AA073862 396 bp mRNA EST 15-FEB-1997
DEFINITION mm97f07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
536389 5' similar to TR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR.
; , mRNA sequence.
ACCESSION AA073862
NID g1595602
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 396)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicinep

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:323325

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 390.

Location/Qualifiers

1. .396

/organism="Mus musculus"

/strain="NIH/Swiss"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3"

/db_xref="taxon:10090"

/clone="536389"

/clone_lib="Stratagene mouse heart (#937316)"

/sex="pooled"

/dev_stage="13 day embryos"

/lab_host="SOLR (kanamycin resistant)"

<1. .>396

93 a 126 c 103 g 74 t

mrna

BASE COUNT

ORIGIN

Query Match 16.2%; Score 328; DB 26; Length 396;

Best Local Similarity 92.5%; Pred. No. 0.00e+00;

Matches 368; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

Db 1 AGCAGCGGGCGCT-GGATTCGCTGGTCATGAAGCAGTGCACCGGACACCAAGAAGT 59

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 476 AGCAGCGGGCGCTTGGATCCCGTGGTCATGAAGCAGTGCACCGGACACCAAGAAGT 535

Db 60 TCCTGTGCTCGCTCTTCGCCCCCTGCTGTCTCGACGACCTAGATGAGACCATCCAGCCGT 119

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 536 TCCTGTGCTCGCTCTTCGCCCCCGCTGCTGCTGATGACCTAGACGAGACCATCCAGCCAT 595

Db 120 GTCACTCGCTCTG-GTGCAGGTGAAGGACCGCTGCGCCCCGGTTCATGTCCGCCCTTCGGCT 178

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 596 GCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTTCATGTCCGCCCTTCGGCT 655

Db 179 TCCCTGGCCAGACATGTGGAGTGGACCGGTTTCCCGCAGGACACACGACCTCTGCATCC 238

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 656 TCCCTGGCCCGACATGCTTGAGTGGACCGGTTTCCCGCAGGACACGACCTTTGCATCC 715

Db 239 CCCTCGCTAGTAGCGACCACTCCTCGCCGGCCACAGAGGAAGTCCCAAGGTGTGTGAAG 298

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 716 CCCTCGCTAGCAGCGACCACTCTCTGCCAGCCACCGAGGAAGTCCAAAGGTATGTGAAG 775

Db 299 CCTGCAAAACCAAGAATGAGGACGACACATCATGGAAACCCCTTTGTAAAAATGACT 358

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 776 CCTGCAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATGATT 835

Db 359 TCGCACTGAAAATCAAAGTGAAGGAGATAACGTACATC 396

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 836 TTGCACCTGAAAAATAAAAGTGAAGGAGAGATAACCTACATC 873

RESULT 30

LOCUS AA024771 414 bp mRNA EST 13-AUG-1996

DEFINITION ze76gl0.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 364962 5', mRNA sequence.

ACCESSION AA024771

NID g1489677

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 414)

REFERENCE Hillier,L., Clark,N., Dubuque,T., Ellliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 288.

Location/Qualifiers

1. .414

/organism="Homo sapiens"

/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGCATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

/db_xref="taxon:9606"

/clone="364962"

/clone_lib="Soares fetal heart NbHH19W"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

<1. .>414

mrna

BASE COUNT 107 a 123 c 95 g 85 t 4 others

ORIGIN

Query Match 16.2%; Score 329; DB 23; Length 414;

Best Local Similarity 97.8%; Pred. No. 0.00e+00;

Matches 360; Conservative 0; Mismatches 3; Indels 5; Gaps 5;

Db 1 TGAAGCAGTGCACCCGGACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCGCTCTGCC 60

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 506 TGAAGCAGTGCACCCGGACACCAAGAAGTCTCTGTGCTCGCTCTTCGCCCGCTCTGCC 565

Db 61 TCGATGACCTAGACGAGACCATCCAGCCATCGCTCTGCGTGCAGGTGAAGGACC 120

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 566 TCGATGACCTAGACGAGACCATCCAGCCATCGCTCTGCGTGCAGGTGAAGGACC 625

Db 121 GCTGCGCCCCGGTCATGTCCGCCCTTCGGNTTCCCTTGGCCCCGAAATGCTTGCAGTCGAC 180

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||


```
QY 626 GCTGCGCCCGTCATGTCCGCCCTTCGGCTTCCCTCGCCCG-ACATGCTTGAGTGGAC 684
Db 181 CGTTTCCCCCAGGACAACGACCTTTTGCATCCCCCTCGCTAGCAGCGACCACTCCTGCCA 240
QY 685 CGTTTCCCCCAGGACAACGACCTTTTGCATCCCCCTCGCTAGCAGCGACCACTCCTGCCA 744
Db 241 GCCACCGAGGAAGCTCCAAAGGTATGTGAAGCCTGCAAAAATAAAAATGATGATGACAAC 300
QY 745 GCCACCGAGGAAGCTCCAAAGGTATGTGAAGCCTGCAAAAATAAAAATGATGATGACAAC 804
Db 301 GACATTAAATGAAACGCTTTGTAAAAATGATTTTGCACCTTGAAAAATTNAAAGTGAAAGGA 360
QY 805 GACAT-AATGGAACGCTTTGTAAAAATGATTTTGCACCT-GAAAT-AAAAGTGAA-GGA 860
Db 361 GATAACCT 368
QY 861 GATAACCT 868

RESULT 31
LOCUS AA705737 329 bp mRNA EST 24-DEC-1997
DEFINITION zf41b02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
ACCESSION AF705737
NID 92715655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 329)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 319.
Location/Qualifiers
1. .329
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."
/db_xref="GDB:1287723"
/db_xref="taxon:9606"
/clone="379467"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
120 a 32 c 59 g 118 t

BASE COUNT 120 a 32 c 59 g 118 t
ORIGIN
```

```
Query Match 15.9%; Score 323; DB 12; Length 329;
Best Local Similarity 99.1%; Pred.No. 0.00e+00;
Matches 326; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 1 TTTAGGTAAACACAGGATGTAAAGTTTATATACAGAATATAATGTTTATCTGAAATATT 60
|||||
Cp 2005 TTTAGGTAAACACAGGATGTAAAGTTTATATACAGAATATAATGTTTATCTGAAATATT 1946
|||||
Db 61 ACAGTGTGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATTAC 120
|||||
Cp 1945 ACAGTGTGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATTAC 1886
|||||
Db 121 AGTAAGCTACAATGGGTTTAATTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCT 180
|||||
Cp 1885 AGTAAGCTACAATGGGTTTAATTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCT 1826
|||||
Db 181 TAAAGTACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTTGGAGAAATCATG 240
|||||
Cp 1825 TAAAGTACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTTGAAGAAATCATG 1766
|||||
Db 241 CTATAGAGATGGTTAATGTGCTTCTAATAATGAAGATATTGTAGCTGGAAATGTGATACA 300
|||||
Cp 1765 CTATAGAAATGGTTAATGTGCTTCTAATAATGAAGATATTGTAGCTGGAAATGTGATACA 1706
|||||
Db 301 TGTAAACAGTTTAAAGTTCCTCCATTGAAGGTA 329
|||||
Cp 1705 TGTAAACAGTTTAAAGTTCCTCCATTGAAGGTA 1677
|||||

RESULT 32
LOCUS H44092 434 bp mRNA EST 31-JUL-1995
DEFINITION YO73a07.r1 Homo sapiens cDNA clone 183540 5'.
ACCESSION H44092
NID g920144
KEYWORDS EST.
SOURCE human clone=183540 library=Soares breast 3NbHBst vector=pT7T3D
(Pharmacia) with a modified polylinker host=DH10B (ampicillin
resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Adult human.
1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGGCCCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
digested with Not I and cloned into the Not I and Eco RI sites of a
modified pT7T3 vector (Pharmacia). Library went through one round
of normalization to a Cot = 20. Library constructed by Bento Soares
and M.Fatima Bonaldo.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 434)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 123
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
1. .434
/organism="Homo sapiens"
```

BASE COUNT 124 a 87 c 72 g 146 t 5 others
ORIGIN

Query Match 15.9%; Score 322; DB 20; Length 434;
Best Local Similarity 93.5%; Pred. NO. 0.00e+00;
Matches 405; Conservative 0; Mismatches 16; Indels 12; Gaps 12;

Db 2 GCTGATTTTCACTTCCCACTGAGGTTGTCTATAACATGCAAAATTCGTTCAATTTCTCTGT 61
|||||
QY 1501 GCTGATTTTCACTTCCCACTGAGGTTGTCTATAACATGCAAAATTCGTTCAATTTCTCTGT 1560
|||||

Db 62 GGCCCAAACCTTGGGTCACAAACCCTGTTGAGATAAAGCTGGCTGTTATCTCAACATCT 121
|||||
QY 1561 GGCCCAAACCTTGGGTCACAAACCCTGTTGAGATAAAGCTGGCTGTTATCTCAACATCT 1620
|||||

Db 122 TCATCAGCTCCAGACTGAGACTCAGTGTCTAAGTCTTACAACAATNCATCATTTTATACC 181
|||||
QY 1621 TCATCAGCTCCAGACTGAGACTCAGTGTCTAAGTCTTACAACAATTCATCAFTTTATACC 1680
|||||

Db 182 TTCAATGGGAACCTTAAACTGTTACATGTATCACATTCAGCTACAAATACATTTTATT 241
|||||
QY 1681 TTCAATGGGAACCTTAAACTGTTACATGTATCACATTCAGCTACAAATACATTTTATT 1740
|||||

Db 242 AGGAAGCACATTAACCAATTTCTATAGGCATGATTTCTCAAGGTAAAGGGCAAAAGGTA 301
|||
QY 1741 AG-AAGCACATTAACCAATTTCTATAG-CATGATTTCTTCAAG-TAAAGG-CAAAAGATA 1796
|||||

Db 302 TTAAATTTTATAATGGACTTGAGGTACTTTTAAAGCCCTTGTTTAAACAATNCCTTACT 361
|||
QY 1797 T-AAATTTTATAATG-ACTTGAG-TACTTT-AAG-CCTTGTTT-AAACAATTCCTTACT 1850
|||||

Db 362 TAACCTTTGGCAATTAACCCCTTGTNGGCTTACCCGGTANTTTACNGGTGGGTTACCT 421
|||||
QY 1851 TAACCTTTGGCAATTAACCCCATTTGTAG-CTTACCTG-TAATATACATAGTAGTTACCT 1908
|||||

Db 422 TTAAAAGTGGTAA 434
|||||
QY 1909 TTAAAAGTTGTAA 1921
|||||

RESULT 33
LOCUS AA335712 321 bp mRNA EST 21-APR-1997
DEFINITION EST40158 Epididymus Homo sapiens cdNA 5' end, mRNA sequence.
ACCESSION AA335712
NID g1987953
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 321)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cdNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 96026280
COMMENT Other_ESTs: THC180057
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavetigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..321
/organism="Homo sapiens"
/note="Organ: epididymis; Vector: pBluescript SK-; Site_1:
EcORI; Site_2: XhoI"
/db_xref="ATCC (inhost):137387"
/db_xref="taxon:9606"
/clone_lib="Epididymus"
/sex="male"
/dev_stage="adult"
<1..>321
mRNA
BASE COUNT 89 a 60 c 63 g 109 t
ORIGIN

Query Match 15.8%; Score 321; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. NO. 0.00e+00;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TGGATAGCTGTTTTCACCTAAAGGAAAGCCACCCGAAATCTTGTAGAAATATTCAAACT 60
|||||
QY 1327 TGGATAGCTGTTTTCACCTAAAGGAAAGCCACCCGAAATCTTGTAGAAATATTCAAACT 1386
|||||

Db 61 AATAAAATCATGAATATTTTATGAAGTTTAAAAATAGCTCACTTTAAAGCTAGTTTGA 120
|||||
QY 1387 AATAAAATCATGAATATTTTATGAAGTTTAAAAATAGCTCACTTTAAAGCTAGTTTGA 1446
|||||

Db 121 ATAGGTGCAACTGTGACTGGGTCTGGTTGGTTGTTGTTGTTTGGTGCAGCTGAT 180
|||||
QY 1447 ATAGGTGCAACTGTGACTGGGTCTGGTTGGTTGTTGTTTGGTGCAGCTGAT 1506
|||||

Db 181 TTTCACTTCCCACTGAGGTTGTCATAACATGCAAAATTCGTTCAATTTTCTGTGGCCCA 240
|||||
QY 1507 TTTCACTTCCCACTGAGGTTGTCATAACATGCAAAATTCGTTCAATTTTCTGTGGCCCA 1566
|||||

Db 241 AACTTGTGGGTCACAAACCCTGTTGAGATAAAGCTGGCTGTTATCTCAACATCTTCATCA 300
|||||
QY 1567 AACTTGTGGGTCACAAACCCTGTTGAGATAAAGCTGGCTGTTATCTCAACATCTTCATCA 1626
|||||

Db 301 GCTCCAGACTGAGACTCAGTG 321
|||||
QY 1627 GCTCCAGACTGAGACTCAGTG 1647
|||||

RESULT 34
LOCUS AA295688 328 bp mRNA EST 18-APR-1997
DEFINITION EST100887 Pancreas tumor I Homo sapiens cdNA 5' end, mRNA sequence.
ACCESSION AA295688
NID g1948033
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 328)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,

Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other_ESTs: THC180057
Contact: Kerlavage, AR
Bioinformatics

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES Location/Qualifiers
source 1..328

/organism="Homo sapiens"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: ECORI; Site_2: XhoI"
/db_xref="AFCC (inhost):191138"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"
/dev_stage="adult"
mRNA <1..>328

BASE COUNT 81 a 68 c 66 g 112 t 1 others
ORIGIN

Query Match 15.8%; Score 320; DB 4; Length 328;
Best Local Similarity 99.4%; Pred. No. 0.00e+00;
Matches 326; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 1 CCTAGCTGCTCCAGCTCTCAGCCTGGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCATCCC 60
|||||
QY 1236 CCTAGCTGCTCCAGCTCTCAGCCTGGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCATCCC 1295

Db 61 CAGCATTTCTGTGAGTTATAAGCCACAGGAGTGGATAGCTGTTTTCACCTAAAGGAAAG 120
|||||
QY 1296 CAGCATTTCTGTGAGTTATAAGCCACAGGAGTGGATAGCTGTTTTCACCTAAAGGAAAG 1355

Db 121 NCCACCCGAATCTTGTAGAAATATTCAAACTAATAAAATCATGAATATTTTATGAAGTT 180
|||||
QY 1356 CCCACCCGAATCTTGTAGAAATATTCAAACTAATAAAATCATGAATATTTTATGAAGTT 1415

Db 181 TAAAAATAGCTCACCTTTAAAGCTAGTTTGAATAGGTGCACTGTGACTTGGGTCTGGTT 240
|||||
QY 1416 TAAAAATAGCTCACCTTTAAAGCTAGTTTGAATAGGTGCACTGTGACTTGGGTCTGGTT 1475

Db 241 GGTGTTGTTTGTGTTTTCAGTCAGCTGATTTTTCACCTCCCACTGAGGTTGTCAATAACA 300
|||||
QY 1476 GGTGTTGTTTGTGTTTTCAGTCAGCTGATTTTTCACCTCCCACTGAGGTTGTCAATAACA 1535

Db 301 TGCAAATTGCTTCAATTTTCTCTGTGG 328
|||||

QY 1536 TGCAAATTGCTTCAATTTT-CTCTGTGG 1562

RESULT 35
LOCUS H15818 469 bp mRNA EST 27-JUN-1995
DEFINITION Y128C05.s1 Homo sapiens cDNA clone 159560 3'.

ACCESSION H15818
NID 9880638
KEYWORDS EST.
SOURCE

human clone=159560 library=Soares breast 3NbHBst vector=pt7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=Promega -21ml3 Rsite1=Not I Rsite2=Eco RI Adult human. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCCTTTTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 20. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 469)

REFERENCE AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 340
Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES Location/Qualifiers
source 1..469
/organism="Homo sapiens"
/clone="159560"

BASE COUNT 144 a 65 c 106 g 150 t 4 others
ORIGIN

Query Match 15.8%; Score 321; DB 20; Length 469;
Best Local Similarity 97.1%; Pred. No. 0.00e+00;
Matches 370; Conservative 0; Mismatches 3; Indels 8; Gaps 8;

Db 1 AGGTAAACACAGGATTTAAAGTTTATATACAAGAATATAATGTTATCTGAAATATTTACA 60
|||||

Cp 2002 AGGTAAACACAGGATGTAAAGTTTATATACAAGAATATAATGTTATCTGAAATATTTACA 1943

Db 61 GTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATACAGG 120
|||||

Cp 1942 GTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATACAGG 1883

Db 121 TAAGCTACAATGGGTTTAATTTCGAAAAGTTAAGTAAGAATGTTTTAAACAAGGCTTAA 180
|||||

Cp 1882 TAAGCTACAATGGGTTTAATTTCGAAAAGTTAAGTAAGAATGTTTTTAAACAAGGCTTAA 1823

Db 181 AGTACTCAAGTCAATTATAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTA 240
|||||

Cp 1822 AGTACTCAAGTCAATTATAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTA 1763

Db 241 TAGAAATGGTTAAATGTGCTTCTTAATAAATGGGAAGTATTGTAGCGTGGGAATGTGATACA 300
|||||

Cp 1762 TAGAAATGGTTAAATGTGCTTCTTAATAAATGGGAAGTATTGTAGCGTGGGAATGTGATACA 1706

Db 301 TGTAACAGTTTAAAGGTTCCCAATGAAGGGGTATAAAATGATGAATTTGTTGAAGGNCITAG 360
|||||

Cp 1705 TGTAACAGTTTAAAG-TTCCCATTTGAAGG-TATAAAATGATGAATTTGTTGTAAG-ACTTAG 1649


```

|||||
QY 1839 ACATTTCTTACTTAACCTTTTGCAAAATTAACCCCATTTGTAGTTACCTGTAATATACATAG 1898
|||||
Db 248 TAGTTTACCTTTAAVAGTTGTAAAAATATTGCTTTAACCAACACTGTAAATATTTTCAGAT 307
|||||
QY 1899 TAGTTTACCTTTTAAAAGTTGTAAAAATATTGCTTTAACCAACACTGTAAATATTTTCAGAT 1958
|||||
Db 308 AAACATTATATTCCYYGTATATAA 330
|||||
QY 1959 AAACATTATATTCTTGTATATAA 1981
|||||

RESULT 38
LOCUS N56835 344 bp mRNA EST 22-FEB-1996
DEFINITION YW83C06.r1 Homo sapiens cDNA clone 258826 5'.
ACCESSION N56835
NID g1200725
KEYWORDS EST.
SOURCE human clone=258826 primer=T7 library=Soares placenta 8to9weeks
2NbHP8to9W vector=pf7T3D (Pharmacia) with a modified polylinker
host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI two
placentae: one from 8 weeks and another from 9 weeks post
conception. 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'-TGTTACCAATCTGAAGTGGGAGCGCGCGATTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pf7T3 vector (Pharmacia). Library
constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 344)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 248
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source Location/Qualifiers
1. .344
/organism="Homo sapiens"
/clone="258826"
<1. .->344
mRNA
BASE COUNT 92 a 70 c 62 g 119 t 1 others
ORIGIN

Query Match 15.4%; Score 313; DB 22; Length 344;
Best Local Similarity 98.3%; Pred. No. 0.00e+00;
Matches 338; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 TGAAGTTTAAAAATAGCTCACTTTAAAGCTAGTTTTGAATAGGTGCAACTGTGACTGGG 60
|||||
QY 1409 TGAAGTTTAAAAATAGCTCACTTTAAAGCTAGTTTTGAATAGGTGCAACTGTGACTGGG 1468
|||||
Db 61 TCTGGTTGGTTGTTGTTGTTGTTGAGTCAGCTGATTTTCACCTCCCACTGAGGTTGT 120
|||||
QY 1469 TCTGGTTGGTTGTTGTTGTTGTTGAGTCAGCTGATTTTCACCTCCCACTGAGGTTGT 1528
|||||
```

```

Db 121 CATAACATGCAAATTGCTTCAATTTTCTGTGGCCCAAACTGTGGGTACAAACCCCTG 180
|||||
QY 1529 CATAACATGCAAATTGCTTCAATTTTCTGTGGCCCAAACTGTGGGTACAAACCCCTG 1588
|||||
Db 181 TTGAGATAAAAGCTGGCTGTTATCTCAACATCTTCATCAGCTCCAGACTGAGACTCAGTGT 240
|||||
QY 1589 TTGAGATAAAAGCTGGCTGTTATCTCAACATCTTCATCAGCTCCAGACTGAGACTCAGTGT 1648
|||||
Db 241 CTAAGTCTTTACAACAATTTCATCATTTTATACCTTCAATGGGAACCTTAAAACTGGTTCCAT 300
|||||
QY 1649 CTAAGTCTTTACAACAATTTCATCATTTTATACCTTCAATGGGAACCTTAAA-CTG-TTACAT 1706
|||||
Db 301 GTATCACAATTCCAGCTACAATACTTCCATTTAATTAGANGCAC 344
|||||
QY 1707 GTATCACA-TTCCAGCTACAATACTTCCATTTA-TTAGAAGCAC 1748
|||||

RESULT 39
LOCUS HUM303B01B 301 bp mRNA EST 14-DEC-1995
DEFINITION Human aorta cDNA 5'-end GEN-303B01, mRNA sequence.
ACCESSION D79675
NID g1180026
KEYWORDS EST; EST(expressed sequence tag); Human aorta.
SOURCE Homo sapiens aorta cDNA to mRNA, clone_lib:human aorta polyA+.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 301)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,
Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H.,
Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y.,
Maekawa,H., Shin,S. and Nakamura,Y.
TITLE Large-scale sequencing project at Otsuka GEN Research Institute
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 301)
AUTHORS Fujiwara,T.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu
Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical
Co.,Ltd, 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,
Japan
COMMENT Submitted (7-Nov-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical CO.,Ltd
463-10 Kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone:0886-65-2888
Fax :0886-37-1035.

FEATURES
source Location/Qualifiers
1. .301
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="human aorta polyA+"
/tissue_type="aorta"
BASE COUNT 106 a 51 c 29 g 115 t
ORIGIN

Query Match 14.8%; Score 299; DB 6; Length 301;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 TACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAGCACATTAACCATTTCT 60
|||||
QY 1702 TACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAGCACATTAACCATTTCT 1761
|||||
Db 61 ATAGCATGATTTCTTCAAGTAAAGGCAAAAGATATAAAATTTTATAATTGACTTGAGTAC 120
|||||
QY 1762 ATAGCATGATTTCTTCAAGTAAAGGCAAAAGATATAAAATTTTATAATTGACTTGAGTAC 1821
|||||
```


Db	1	CTGAGACTCAGTGTCTAAGTCTTACAACAATTTCATCATCTTTATATACCTTCAATGGGAAC	60
QY	1635	CTGAGACTCAGTGTCTAAGTCTTACAACAATTTCATCATCTTTATATACCTTCAATGGGAAC	1694
Db	61	AAACTGTTACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAGGCACATTAA	120
QY	1695	AAACTGTTACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAGGCACATTAA	1754
Db	121	CATTCTATAGCATGATTTCTCAAGTAAAAGGCAAAAGATATAAATTTTATAATTGACT	180
QY	1755	CATTCTATAGCATGATTTCTCAAGTAAAAGGCAAAAGATATAAATTTTATAATTGACT	1814
Db	181	TGAGTACTTTAAGCCTTGTTAAACATTTCTTACTTAACTTTTGCAA-TTAM-CCCA	238
QY	1815	TGAGTACTTTAAGCCTTGTTAAACATTTCTTACTTAACTTTTGCAAATTTAAACCCAT	1874
Db	239	GTAGCTTACCTGTAATATACATAGTAGTTTACCTTTTAARVGTGTAAAAATATTKC	298
QY	1875	GTAGCTTACCTGTAATATACATAGTAGTTTACCTTTTAARVGTGTAAAAATATTGC	1934
Db	299	MCCAACACTGTAAAT	313
QY	1935	ACCAACACTGTAAAT	1949
RESULT	42	HUM347H05B	303 bp mRNA EST 21-MAY-1996
LOCUS		Human aorta cDNA 5'-end GEN-347H05,	mRNA sequence.
DEFINITION		D62999	
ACCESSION		g966773	
NID		EST; EST(expressed sequence tag); Human aorta; similar to none(May	
KEYWORDS		29,1995).	
SOURCE		Homo sapiens cDNA to mRNA, clone_lib:Clontech human aorta polyA+	
ORGANISM		mRNA (#6572).	
REFERENCE		Homo sapiens	
AUTHORS		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;	
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;	
		Homo.	
		1 (bases 1 to 303)	
		Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,	
		Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H.,	
		Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y.,	
		Maekawa,H., Shin,S. and Nakamura,Y.	
		Unpublished(303)	
		Unpublished (1995)	
		2 (bases 1 to 303)	
		Fujiwara,T.	
		Direct Submission	
		Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu	
		Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical	
		Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,	
		Japan (Tel:0886-65-2888, Fax:0886-37-1035)	
		Submitted (30-May-1995) to DDBJ by:	
		Tsutomu Fujiwara	
		Otsuka GEN Research Institute	
		Otsuka Pharmaceutical Co.,Ltd	
		463-10 kagasuno Kawauchi-cho	
		Tokushima, Tokushima	
		771-01	
		Japan	
		Phone: 0886-65-2888	
		Fax : 0886-37-1035.	
FEATURES		Location/Qualifiers	
source		1. .303	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"	
BASE COUNT	105 a	52 c	32 g 111 t 3 others
ORIGIN			
Query Match	14.7%	Score 298;	DB 6; Length 303;
Best Local Similarity	99.0%	Pred. No. 0.00e+00;	
Matches	297;	Conservative	2; Mismatches 1; Indels 0; Gaps 0;

Db	4	ACTCAGTGTCTAAGTCTTACAACAATTTCATCATCTTTATATACCTTCAATGGGAAC	63
QY	1640	ACTCAGTGTCTAAGTCTTACAACAATTTCATCATCTTTATATACCTTCAATGGGAAC	1699
Db	64	GTTACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAGGCACATTAA	123
QY	1700	GTTACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAGGCACATTAA	1759
Db	124	CTATAGCATGATTTCTTCAAGTAAAAGGCAAAAGATATAAATTTTATAATTGAGT	183
QY	1760	CTATAGCATGATTTCTTCAAGTAAAAGGCAAAAGATATAAATTTTATAATTGAGT	1819
Db	184	ACTTTAAGCCTTGTTTAAACATTTCTTACTTAACTTTTGCAAAATTAACCCATTG	243
QY	1820	ACTTTAAGCCTTGTTTAAACATTTCTTACTTAACTTTTGCAAAATTAACCCATTG	1879
Db	244	TTACCTGTAAATATACATAGTAGTTTACCTTTTAARAGTTGTAAAAGTATTGCTTT	303
QY	1880	TTACCTGTAAATATACATAGTAGTTTACCTTTTAARAGTTGTAAAAGTATTGCTTT	1939
RESULT	43	AA347786	300 bp mRNA EST 21-APR-1997
LOCUS		EST54420 Fetal heart II Homo sapiens cDNA 3'	end, mRNA sequence.
DEFINITION		AA347786	
ACCESSION		g2000261	
NID		EST.	
KEYWORDS		human.	
SOURCE		Homo sapiens	
ORGANISM		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;	
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;	
		Homo.	
		1 (bases 1 to 300)	
		Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,	
		Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,	
		White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,	
		Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,	
		Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,	
		Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,	
		Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,	
		Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,	
		Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,	
		Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,	
		Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,	
		Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,	
		He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,	
		Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,	
		Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,	
		Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,	
		Fraser,C.M. and Venter,J.C.	
		Initial assessment of human gene diversity and expression patterns	
		based upon 83 million nucleotides of cDNA sequence	
		Nature 377 (6547 Suppl), 3-174 (1995)	
		96026280	
		Other_ESTs: EST113252 THC180057	
		Contact: Kerlavage, AR	
		Bioinformatics	
		The Institute for Genomic Research	
		9712 Medical Center Drive, Rockville, MD 20850 USA	
		Tel: 3018699056	
		Fax: 3018699423	
		Email: arkerlavet@tigr.org	
		For clone availability, additional sequence and expression	
		information related to this EST, please check the TIGR Human Gene	
		Index (http://www.tigr.org/tldb/hgi/hgi.html)	
		Seq primer: M13-21.	
		Location/Qualifiers	
		1. .300	
		/organism="Homo sapiens"	
		/note="Organ: heart; Vector: pBluescript SK-; Site_1:	
		EcoRI; Site_2: XhoI"	
		/db_xref="ATCC (inhost):149068"	
FEATURES			
source			

Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source
1. .302
/organism="Homo sapiens"
/note="Organ: embryo; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
/db_xref="AtCC (inhost):132379"
/db_xref="taxon:9606"
/clone_lib="Embryo, 6 week I"
/dev_stage="embryo, 6 wks"
<1. .>302

BASE COUNT 99 a 57 c 82 g 56 t 8 others
ORIGIN

Query Match 14.4%; Score 292; DB 4; Length 302;
Best Local Similarity 97.0%; Pred. No. 0.00e+00;
Matches 293; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 1 AACGACATAATGGAAACGGCTTGTAAAAATNATTTNCACTGAAAAATAAAAGTGAAGGAG 60
|||||
QY 802 AACGACATAATGGAAACGGCTTGTAAAAATGATTTTGCACCTGAAAAATAAAAGTGAAGGAG 861
Db 61 ATAACCTACATCAACCGGAGATACCAAAATCATCCTGGAGACCAAGAGCAAGACNNTTTAC 120
|||||
QY 862 ATAACCTACATCAACCGGAGATACCAAAATCATCCTGGAGACCAAGAGCAAGACATTTAC 921
Db 121 AAGCTGAACGGTGTGCCGAAAGGGACCTGAAGAAATCGTGCTGTGGCTCAAAGACAGC 180
|||||
QY 922 AAGCTGAACGGTGTGCCGAAAGGGACCTGAAGAAATCGTGCTGTGGCTCAAAGACAGC 981
Db 181 TTGCAGTGCACCTNTNAGGAGATGAACGACATCAACGNGGCCCTATCTGGTCATGGGACAG 240
|||||
QY 982 TTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGCGCCTATCTGGTCATGGGACAG 1041
Db 241 AANCAGGGTTGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGAAAGGGCAGAGAGAG 300
|||||
QY 1042 AAACAGGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGAAAGGGCAGAGAGAG 1101

Db 301 TT 302
||
QY 1102 TT 1103

RESULT 46 HUM238C04B 293 bp mRNA EST 21-MAY-1996
LOCUS Human aorta cDNA 5'-end GEN-238C04, mRNA sequence.
DEFINITION D62097
ACCESSION 9965873
NID
KEYWORDS EST; EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).
SOURCE Homo sapiens cDNA to mRNA, clone_lib:Clontech human aorta polyA+ mRNA (#6572).
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 293)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.
TITLE Unpublished(303)
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 293)
AUTHORS Fujiwara,T.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,

COMMENT
Japan (Tel:0886-65-2888, Fax:0886-37-1035)
Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co.,Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.

FEATURES
source
1. .293
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"

BASE COUNT 102 a 51 c 26 g 109 t 5 others
ORIGIN

Query Match 14.3%; Score 290; DB 6; Length 293;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 287; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 1 TCACATTCCAGCTACAATACTTCCATTATTAGAGCACATTAAACCATTTCTATAGCATG 60
|||||
QY 1710 TCACATTCCAGCTACAATACTTCCATTATTAGAGCACATTAAACCATTTCTATAGCATG 1769
Db 61 ATTTCTTCAAGTAAAGGCAAGAGATATAAMTYYYATAATCACTTGAGTACTTTAAGCC 120
|||||
QY 1770 ATTTCTTCAAGTAAAGGCAAGAGATATAAATTTTATAATTGACTTGAGTACTTTAAGCC 1829
Db 121 TTGTTTAAACACATTCTTACTTAACCTTTTGCAAATTAACCCATTGTAGCTTACCTGTAA 180
|||||
QY 1830 TTGTTTAAACACATTCTTACTTAACCTTTTGCAAATTAACCCATTGTAGCTTACCTGTAA 1889
Db 181 TATACATAGTAGTTTACCCTTAAAGTTGTAAAAATATTTGCTTTAACCAACACTGTAAAT 240
|||||
QY 1890 TATACATAGTAGTTTACCCTTAAAGTTGTAAAAATATTTGCTTTAACCAACACTGTAAAT 1949
Db 241 ATTTCAGATAAACATTATATCTTGTATATAANCCTTTACATCCTGTTTACCT 293
|||||
QY 1950 ATTTCAGATAAACATTATATCTTGTATATAAACTTTACATCCTGTTTACCT 2002

RESULT 47 HUM238C05B 293 bp mRNA EST 21-MAY-1996
LOCUS Human aorta cDNA 5'-end GEN-238C05, mRNA sequence.
DEFINITION D62098
ACCESSION 9965874
NID
KEYWORDS EST; EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).
SOURCE Homo sapiens cDNA to mRNA, clone_lib:Clontech human aorta polyA+ mRNA (#6572).
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 293)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.
TITLE Unpublished(303)
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 293)
AUTHORS Fujiwara,T.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,

COMMENT
Japan (Tel:0886-65-2888, Fax:0886-37-1035)
Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara

COMMENT Japan (Tel:0886-65-2888, Fax:0886-37-1035)
Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co.,Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.
Location/Qualifiers
source
1. .299
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
BASE COUNT 104 a 51 c 28 g 112 t 4 others
ORIGIN

Query Match 14.2%; Score 287; DB 6; Length 299;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 295; Conservative 4; Mismatches 0; Indels 2; Gaps 2;

Db 1 TACATGTATCACATTCAGCTACAACTTCCATTTATTAGAGCACATTAAACCATTTCT 60
|||||
QY 1702 TACATGTATCACATTCAGCTACAACTTCCATTTATTAGAGCACATTAAACCATTTCT 1761
|||||

Db 61 ATAGCATGATTTCCTCAAGTAAAGGCAAGAGATATAAATTTTATAATTGACTTGAGTAC 120
|||||
QY 1762 ATAGCATGATTTCCTCAAGTAAAGGCAAGAGATATAAATTTTATAATTGACTTGAGTAC 1821
|||||

Db 121 TTAAAGCCTTGTTTAAACATTTCTTACTTAACTTTTGCAAAATTAAACCCATTGTAGCTT 180
|||||
QY 1822 TTAAAGCCTTGTTTAAACATTTCTTACTTAACTTTTGCAAAATTAAACCCATTGTAGCTT 1881
|||||

Db 181 ACCYGTAATATACATAGTAGTTACCTTTAAAAGTTGTAAAA-TATTGCTTTAAACCAMCA 239
|||:|||||
QY 1882 ACCTGTAATATACATAGTAGTTACCTTTAAAAGTTGTAAAAATATTGCTTTAAACCAACA 1941
|||||

Db 240 CTGTAAATATTTTCAGATAAACATATATATCTTGTATATAAACYT-ACATCCTGTTTACC 298
|||||
QY 1942 CTGTAAATATTTTCAGATAAACATATATATCTTGTATATAAACATTACATCCTGTTTACC 2001
|||||

Db 299 T 299
QY 2002 T 2002

RESULT 50
LOCUS HUM289F09B 287 bp mRNA EST 21-MAY-1996
DEFINITION Human aorta cDNA 5'-end GEN-289F09, mRNA sequence.
ACCESSION D62477
NID 9966253
KEYWORDS EST; EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).
SOURCE Homo sapiens cDNA to mRNA, clone_lib:Clontech human aorta polyA+ mRNA (#6572).
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 287)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.
Unpublished(303)
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 287)
AUTHORS Fujiwara,T.
TITLE Direct Submission
Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical

COMMENT Co.,Ltd: 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,
Japan (Tel:0886-65-2888, Fax:0886-37-1035)
Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co.,Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.
Location/Qualifiers
source
1. .287
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
BASE COUNT 99 a 50 c 33 g 105 t
ORIGIN

Query Match 14.1%; Score 285; DB 6; Length 287;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 TGAGACTCAGTGTCTAAGTCTTACAACAATTTCATCATTTTATACCTTCAATGGGAACCTTA 60
|||||
QY 1636 TGAGACTCAGTGTCTAAGTCTTACAACAATTTCATCATTTTATACCTTCAATGGGAACCTTA 1695
|||||

Db 61 AACTGTTACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAAGCACATTAAACC 120
|||||
QY 1696 AACTGTTACATGTATCACATTCAGCTACAATACTTCCATTTATTAGAAGCACATTAAACC 1755
|||||

Db 121 ATTTCTATAGCATGATTTCTTCAAGTAAAGGCAAGAGATATAAATTTTATAATTGACTT 180
|||||
QY 1756 ATTTCTATAGCATGATTTCTTCAAGTAAAGGCAAGAGATATAAATTTTATAATTGACTT 1815
|||||

Db 181 GAGTACTTTAAGCCTTGTTTAAACATTTCTTACTTAACTTTTGCAAAATTAAACCCATTG 240
|||||
QY 1816 GAGTACTTTAAGCCTTGTTTAAACATTTCTTACTTAACTTTTGCAAAATTAAACCCATTG 1875
|||||

Db 241 TAGCTTACCTGTAATATACATAGTAGTTTACCTTTTAAAGGTTGTAAA 287
|||||
QY 1876 TAGCTTACCTGTAATATACATAGTAGTTTACCTTTTAAAGTTGTAAA 1922
|||||

RESULT 51
LOCUS HUM343G04B 281 bp mRNA EST 15-DEC-1995
DEFINITION Human aorta cDNA 5'-end GEN-343G04, mRNA sequence.
ACCESSION D79867
NID g1180218
KEYWORDS EST; EST(expressed sequence tag); Human aorta.
SOURCE Homo sapiens aorta cDNA to mRNA, clone_lib:human aorta polyA+.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 281)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.
Large-scale sequencing project at Otsuka GEN Research Institute
Unpublished (1995)
REFERENCE 2 (bases 1 to 281)
AUTHORS Fujiwara,T.
TITLE Direct Submission
Submitted (07-NOV-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical Co.,Ltd, 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01, Japan
Submitted (7-Nov-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute

Otsuka Pharmaceutical CO.,Ltd
463-10 Kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone:0886-65-2888
Fax :0886-37-1035.

FEATURES
source
Location/Qualifiers
1..281

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="human aorta polyA+"
/tissue_type="aorta"

BASE COUNT 101 a 45 c 24 g 110 t 1 others
ORIGIN

Query Match 13.8%; Score 279; DB 6; Length 281;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 279; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 TACAATACTTCCATTATTATTAGAGCACATTAAACCATTTCTATAGCATGATTTCCTCAAGT 60
|||||
QY 1722 TACATACTTCCATTATTATTAGAGCACATTAAACCATTTCTATAGCATGATTTCCTCAAGT 1781
Db 61 ARAAGGCAAAAGATATAAATTTTATAATTGACTTGACTTGAGTACTTTAAGCCTTGTTTAAACA 120
|||||
QY 1782 ARAAGGCAAAAGATATAAATTTTATAATTGACTTGACTTGAGTACTTTAAGCCTTGTTTAAACA 1841
Db 121 TTCTTACTTAACTTTTGCAATTAACCCCATTTGTAGCTTACCTGTAATATACATAGTAG 180
|||||
QY 1842 TTCTTACTTAACTTTTGCAATTAACCCCATTTGTAGCTTACCTGTAATATACATAGTAG 1901
Db 181 TTTACCTTTAAAGTTGTAAATATATTGCTTTAACCAACACTGTAAATATTTTCAGATAAA 240
|||||
QY 1902 TTTACCTTTAAAGTTGTAAATATATTGCTTTAACCAACACTGTAAATATTTTCAGATAAA 1961
Db 241 CATTATATTCTTKTATATATAAACTTTACATCCTTTTACCT 281
|||||
QY 1962 CATTATATTCTTGTATATAAACTTTACATCCTGTTTACCT 2002

RESULT 52
LOCUS HUM233G01B 273 bp mRNA EST 21-MAY-1996
DEFINITION Human aorta cDNA 5'-end GEN-233G01, mRNA sequence.
ACCESSION D62031
NID 9965807
KEYWORDS EST; EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).
SOURCE Homo sapiens cDNA to mRNA, clone_lib:Clontech human aorta polyA+ mRNA (#6572).

ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 273)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.
Unpublished(303)
Unpublished (1995)
REFERENCE 2 (bases 1 to 273)
AUTHORS Fujiwara,T.
TITLE Direct Submission

JOURNAL Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01, Japan (Tel:0886-65-2888, Fax:0886-37-1035)
REFERENCE Submitted (30-May-1995) to DDBJ by:
AUTHORS Tsutomu Fujiwara
TITLE Otsuka GEN Research Institute
JOURNAL Otsuka Pharmaceutical Co.,Ltd
COMMENT 463-10 kagasuno Kawauchi-cho

Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.

FEATURES
source
Location/Qualifiers
1..273

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
100 a 43 c 25 g 99 t 6 others

Query Match 13.5%; Score 273; DB 6; Length 273;
Best Local Similarity 97.8%; Pred. No. 0.00e+00;
Matches 267; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Db 1 ATCACATTCAGCTACAATACTTCCATTATTAGAGCACATTAAACCATTTCTATAGCAT 60
|||||
QY 1709 ATCACATTCAGCTACAATACTTCCATTATTAGAGCACATTAAACCATTTCTATAGCAT 1768
Db 61 GATTTCTTCAAGTAAAGGCAAAAGATATAAATYYBATAATTVMMTTGAGTACTTTAAGC 120
|||||
QY 1769 GATTTCTTCAAGTAAAGGCAAAAGATATAAATTTTATAATTGACTTGAGTACTTTAAGC 1828
Db 121 CTTGTTTAAACATTTCTTACTTAACTTTGCAAAATTAACCCATTGTAGCTTACCTGTA 180
|||||
QY 1829 CTTGTTTAAACATTTCTTACTTAACTTTGCAAAATTAACCCATTGTAGCTTACCTGTA 1888
Db 181 ATATACATAGTAGTTTACCTTTTAAAGTTGTAAAAATATTGCTTTAACCAACACTGTAAA 240
|||||
QY 1889 ATATACATAGTAGTTTACCTTTTAAAGTTGTAAAAATATTGCTTTAACCAACACTGTAAA 1948
Db 241 TATTTTCAGATAAAACATTATATTCTTGTATATAA 273
|||||
QY 1949 TATTTTCAGATAAAACATTATATTCTTGTATATAA 1981

RESULT 53
LOCUS AA371169 276 bp mRNA EST 21-APR-1997
DEFINITION EST83193 Prostate gland I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA371169
NID g2023731
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 276)

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Melssner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

TITLE

JOURNAL

MEDLINE

COMMENT

Nature 377 (6547 Suppl), 3-174 (1995)
96026280
Other_ESTs: THC180057
Contact: Kerlavage, AR

Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source
1. .276
/organism="Homo sapiens"
/note="Organ: prostate; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI"
/db_xref="ATCC (inhost):175811"
/db_xref="taxon:9606"
/clone_lib="Prostate gland I"
/sex="male"
/dev_stage="adult, 21 yrs"
<1. .>276

BASE COUNT 52 a 92 c 72 g 60 t
ORIGIN

Query Match 13.5%; Score 274; DB 5; Length 276;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 275; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GTGATCACCTCGGTGAAGCGGTGGCAGAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGC 60
|||||
QY 1060 GTGATCACCTCGGTGAAGCGGTGGCAGAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGC 1119
|||||

Db 61 AGCATCCGCAAGCTGCAGTGCTAGTCCCGGCATCCTTATGGCTCCGACAGGCGCTGCTCCA 120
|||||
QY 1120 AGCATCCGCAAGCTGCAGTGCTAGTCCCGGCATCCTGATGGCTCCGACAGGCGCTGCTCCA 1179
|||||

Db 121 GAGCAGGGCTGACCATTTCTGCTCGGGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTA 180
|||||
QY 1180 GAGCAGGGCTGACCATTTCTGCTCGGGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTA 1239
|||||

Db 181 GCTGCTCCAGTCTCAGCCTGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCATCCCCCAGC 240
|||||
QY 1240 GCTGCTCCAGTCTCAGCCTGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCATCCCCCAGC 1299
|||||

Db 241 ATTTCTGAGTTATAAGGCCACAGGAGTGGATAGCT 276
|||||
QY 1300 ATTTCTGAGTTATAAGGCCACAGGAGTGGATAGCT 1335
|||||

RESULT 54
LOCUS AA339113 282 bp mRNA EST 21-APR-1997
DEFINITION EST44185 Fetal brain I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA339113
NID g1991371
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (bases 1 to 282)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palanco,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,

He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
Other_ESTs: THC180057
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source
1. .282
/organism="Homo sapiens"
/note="Organ: brain; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI"
/db_xref="ATCC (inhost):140742"
/db_xref="taxon:9606"
/clone_lib="Fetal brain I"
/sex="female"
/dev_stage="fetus, 24 wks"
<1. .>282

BASE COUNT 69 a 70 c 56 g 82 t 5 others
ORIGIN

Query Match 13.5%; Score 273; DB 4; Length 282;
Best Local Similarity 97.5%; Pred. No. 0.00e+00;
Matches 275; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 1 TTCTGCTCCGGGATCTCAGCTCCCGTTCCCCAAGCACAGTCTCAGTCTCAGTCTCAG 60
|||||
QY 1196 TTCTGCTCCGGGATCTCAGCTCCCGTTCCCCAAGCACAGTCTCAGTCTCAGTCTCAG 1255
|||||

Db 61 CCTGGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTGAGTTATAA 120
|||||
QY 1256 CCTGGGCAGCTTCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTGAGTTATAA 1315
|||||

Db 121 GGCCACAGGAGTNGNTAGTGTGTTTTCACCTAAAGGAAAGCCACCCGAATCTGTAGNA 180
|||||
QY 1316 GGCCACAGGAGTGGATAGTGTGTTTTCACCTAAAGGAAAGCCACCCGAATCTGTAGAA 1375
|||||

Db 181 ATATTCAAACATAAAATCATGAATATTTTATGAAGGTTAAATAAGTCACTCTTAA 240
|||||
QY 1376 ATATTCAAACATAAAATCATGAATATTTTATGAAGTTTAAATAAGTCACTCTTAA 1435
|||||

Db 241 GCTAGTTTTGANTAGGTGCAACTNTGACTTGGTCTGGTTGG 282
|||||
QY 1436 GCTAGTTTTGANTAGGTGCAACTGTGACTTGGTCTGGTTGG 1477
|||||

RESULT 55
LOCUS C16543 283 bp mRNA EST 04-SEP-1996
DEFINITION Human aorta cDNA 5'-end GEN-347D02, mRNA sequence.
ACCESSION C16543
NID g1571250
KEYWORDS EST; EST(expressed sequence tag); Human aorta.
SOURCE Homo sapiens aorta cDNA to mRNA, clone:347D02.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.

REFERENCE 1 (sites)
AUTHORS Fujiwara,T., Hirano,H., Hishigaki,H., Horie,M., Kawai,A., Kuga,Y.,

Kyushiki,H., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y.,
Shinomiya,H., Suzuki,M., Takaichi,A., Takeda,S., Watanabe,T.,
Maekawa,H., Nakamura,Y. and Takahashi,E.

TITLE
JOURNAL
REFERENCE 2 (bases 1 to 283)
AUTHORS Fujiwara,T.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-1996) to the DDBJ/EMBL/GenBank databases. Tsutomu Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01, Japan (Tel:+81-886-65-2888, Fax:+81-886-37-1035)

FEATURES
source
1. .283
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="347D02"
/tissue_type="aorta"

BASE COUNT 101 a 46 c 26 g 110 t
ORIGIN
Query Match 13.3%; Score 269; DB 6; Length 283;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 281; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Db 1 TACAATACCTCCATTATTAGAAAGCACATTAAACCATTTCTATAGCATGATTTCTCAAGT 60
|||||
QY 1722 TACAATACCTCCATTATTAGAAAGCACATTAAACCATTTCTATAGCATGATTTCTCAAGT 1781
Db 61 AAAAGGCAAAAGATATAAATTTTATAATTGACTTGAGTACTTTAAGCCCTGTTTAAACA 120
|||||
QY 1782 AAAAGGCAAAAGATATAAATTTTATAATTGACTTGAGTACTTTAAGCCCTGTTTAAACA 1841
Db 121 TTTCTTACTTAACCTTTTGCAAAATTAAACCCATTGTAGCTTACCTGTAATATACATAGTAG 180
|||||
QY 1842 TTTCTTACTTAACCTTTTGCAAAATTAAACCCATTGTAGCTTACCTGTAATATACATAGTAG 1901
Db 181 TTTACCTTTAAAAGTTGTAAAAATATTGCTTTAACCAACCACTGTAATATTTTCAGATAA 240
|||||
QY 1902 TTTACCTTTAAAAGTTGTAAAAATATTGCTTTAACCAAC-ACGTGAATATTTTCAGATAA 1960
Db 241 ACATTATATCTGTATATATAAACCTTTTACATCCTCTGTTTACCT 283
|||||
QY 1961 ACATTATATCTGTATATATAAACCTTT-ACATCCTGTTTACCT 2002

RESULT 56
LOCUS AA176165 262 bp mRNA EST 30-DEC-1996
DEFINITION zp23hl2.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone 610343 3', mRNA sequence.
ACCESSION AA176165
NID gl757287
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 262)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13.fwd. from Amersham
High quality sequence stop: 161.

FEATURES
source
1. .262
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. NT2 cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24 hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
/db_xref="taxon:9606"
/clone="610343"
/clone_lib="Stratagene neuroepithelium (#937231)"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
complement(<1.>262)

BASE COUNT 100 a 27 c 40 g 91 t 4 others
ORIGIN

Query Match 12.6%; Score 256; DB 25; Length 262;
Best Local Similarity 98.1%; Pred. No. 0.00e+00;
Matches 257; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 GGTAACAACAGGATGTAAGTTTATATACAAGAATATAATGTTTATCTGAAATATTACAG 60
|||||
Cp 2001 GGTAACAACAGGATGTAAGTTTATATACAAGAATATAATGTTTATCTGAAATATTACAG 1942
Db 61 TGTTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGGT 120
|||||
Cp 1941 TGTTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGGT 1882
Db 121 AAGCTACAATGGGTTTAATTGCAAAAGTTTAAGTAAGAAATGTTTAAACAAGGCTTAAA 180
|||||
Cp 1881 AAGCTACAATGGGTTTAATTGCAAAAGTTTAAGTAAGAAATGTTTAAACAAGGCTTAAA 1822
Db 181 GTACTCAAGTCAATTATAAAAATTTATANCCTTTNGCCTTTTACTTGAAGAAATCATGCTAT 240
|||||
Cp 1821 GTACTCAAGTCAATTATAAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTAT 1762
Db 241 AGAAATGGNTAATGTCCTCTA 262
|||||
Cp 1761 AGAAATGGTTAATGTCCTTCTA 1740

RESULT 57
LOCUS HUM286A04B 263 bp mRNA EST 14-DEC-1995
DEFINITION Human aorta cDNA 5'-end GEN-286A04, mRNA sequence.
ACCESSION D79597
NID g1179948
KEYWORDS EST; EST(expressed sequence tag); Human aorta.
SOURCE Homo sapiens aorta cDNA to mRNA, clone_lib:human aorta polyA+.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 263)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.
TITLE Large-scale sequencing project at Otsuka GEN Research Institute
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 263)
AUTHORS Fujiwara,T.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical Co.,Ltd, 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01, Japan
COMMENT Submitted (7-Nov-1995) to DDBJ by:
Tsutomu Fujiwara

Otsuka GEN Research Institute
Otsuka Pharmaceutical CO.,Ltd
463-10 Kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone:0886-65-2888
Fax :0886-37-1035.

FEATURES
source
Location/Qualifiers
1. .263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="human aorta polyA+"
/tissue_type="aorta"
94 a 45 c 28 g 93 t 3 others
ORIGIN

Query Match 12.6%; Score 256; DB 6; Length 263;
Best Local Similarity 98.5%; Pred. NO. 0.00e+00;
Matches 259; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

Db 1 TCAATGGGAACCTTAACTGTTACATGTATCACATTCAGCTACAATACTTCCATTATTA 60
|||||
QY 1682 TCAATGGGAACCTTAACTGTTACATGTATCACATTCAGCTACAATACTTCCATTATTA 1741
|||||

Db 61 GAAGCACATTAACTTCTATAGCATGATTTCTTCAAGTAAAGGCAAAAGATATAAAT 120
|||||
QY 1742 GAAGCACATTAACTTCTATAGCATGATTTCTTCAAGTAAAGGCAAAAGATATAAAT 1801
|||||

Db 121 YTTATAATTSACTTGAGTACTTTAAGCCTTGTTTAAACATTTCTTACTTAACTTTGCA 180
:|||||
QY 1802 TTTATAATTGACTTGAGTACTTTAAGCCTTGTTTAAACATTTCTTACTTAACTTTGCA 1861
|||||

Db 181 AATTAAACCCATTGTAGCTTACCYGTAAATATACATAGTAGTTTACCTTTAAAGTTGTA 240
|||||
QY 1862 AATTAAACCCATTGTAGCTTACCTGTAAATATACATAGTAGTTTACCTTTAAAG-TTGTA 1920
|||||

Db 241 AAAATATTGCTTTTAAACCAACT 263
|||||
QY 1921 AAAATATTGCTTTTAAACCAACT 1943
|||||

RESULT 58
LOCUS AA082155 287 bp mRNA EST 01-FEB-1997
DEFINITION ze88f06.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 366083 5', mRNA sequence.
ACCESSION AA082155
NID g1624408
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 287)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
WashU-Merck EST Project
Unpublished (1995)
TITLE This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
JOURNAL Insert Length: 1171 Std Error: 0.00
COMMENT Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 262.

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

Location/Qualifiers
1. .287
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."
/db_xref="taxon:9606"
/clone="366083"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1. .>287
62 a 84 c 85 g 51 t 5 others
ORIGIN

mRNA
BASE COUNT 62 a 84 c 85 g 51 t 5 others
ORIGIN

Query Match 12.6%; Score 255; DB 25; Length 287;
Best Local Similarity 97.5%; Pred. NO. 0.00e+00;
Matches 273; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

Db 1 GGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGGCCCTATC 60
|||||
QY 968 GGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGGCCCTATC 1027
|||||

Db 61 TGGTCATGGGACAGAAACAGGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGA 120
|||||
QY 1028 TGGTCATGGGACAGAAACAGGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGA 1087
|||||

Db 121 AGGGGCAGAGAGAGTTCAAGCGCATCTNCCGGNNGNATCCGCAAGCTGCAGTGTCTAGTCC 180
|||||
QY 1088 AGGGGCAGAGAGAGTTCAAGCGCATCTCCCG-CAGCATCCGCAAGCTGCAGTGTCTAGTCC 1146
|||||

Db 181 CGGCATCCTGATGGCTCCGACAG-CCTGTCTCCAGAGGACGGCTGACCATTTCGTCTCCGG 239
|||||
QY 1147 CGGCATCCTGATGGCTCCGACAGGCCTGTCTCCAGAGGACGGCTGACCATTTCGTCTCCGG 1206
|||||

Db 240 GATCTCAGCTCCCGTTCCC-AAGCACACTCCTAGTGTCTC 278
|||||
QY 1207 GATCTCAGCTCCCGTTCCCAAGCACACTCCTAGTGTCTC 1246
|||||

RESULT 59
LOCUS AA260087 394 bp mRNA EST 18-MAR-1997
DEFINITION va52g08.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 735038 5', mRNA sequence.
ACCESSION AA260087
NID g1896590
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 394)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
TITLE This WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project

Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:452086

Seq primer: -28ml3 rev2 ET from Amersham.

FEATURES
source

Location/Qualifiers
1..394
/organism="Mus musculus"
/strain="C57BL/6J"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGAGCGCGCTTATTTTCTTTTCTTTT 3'],
on total mouse RNA [provided by Minoru Ko, Wayne State
Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:10090"
/clone="735038"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
<1..>394

BASE COUNT 104 a 96 c 118 g 76 t
ORIGIN

Query Match 12.6%; Score 255; DB 26; Length 394;
Best Local Similarity 94.1%; Pred. No. 0.00e+00;
Matches 272; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Db 1 AAGGAGATAACGTACATCAACAGAGACACCAAGATCATCTGGAGACAAAGCAAGACC 60
|||||
QY 856 AAGGAGATAACCTACATCAACCGAGATACCAAAATCATCTCGAGACCAAGCAAGACC 915
|||||
Db 61 ATTTACAAGCTGAACGGCGGTGCCGAAAGGGACCTGAAGAAATCCGTGCTGGCTCAA 120
|||||
QY 916 ATTTACAAGCTGAACGGTGTGCCGAAAGGGACCTGAAGAAATCCGTGCTGGCTCAA 975
|||||
Db 121 GACAGCCTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGCTCCGTATCTGGTCATG 180
|||||
QY 976 GACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGCCCTATCTGGTCATG 1035
|||||
Db 181 GGACAGAAGCAGGGCGGAGCTGGTGATCACCTCCGTGAACGGTGGCAGAGGGCCAG 240
|||||
QY 1036 GGACAGAAACAGGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGAGGGGCAG 1095
|||||
Db 241 AGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAATGCTAGT 289
|||||
QY 1096 AGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTAGT 1144
|||||

RESULT 60
LOCUS HUM224H07B 266 bp mRNA EST 21-MAY-1996
DEFINITION Human aorta cDNA 5'-end GEN-224H07, mRNA sequence.
ACCESSION D61938
NID 9965714
KEYWORDS EST; EST(expressed sequence tag); Human aorta; similar to none(May
29,1995).
SOURCE Homo sapiens cDNA to mRNA, clone_lib:Clontech human aorta polyA+
mRNA (#6572).
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 266)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,

Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H.,
Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y.,
Maekawa,H., Shin,S. and Nakamura,Y.
Unpublished(303)
Unpublished (1995)
2 (bases 1 to 266)
Fujiwara,T.
Direct Submission
Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu
Fujiwara, Otsuka GEN Research Institute,Otsuka Pharmaceutical
Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,
Japan (Tel:0886-65-2888, Fax:0886-37-1035)
Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co.,Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.

COMMENT

FEATURES
source

Location/Qualifiers
1..266
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
BASE COUNT 92 a 45 c 26 g 98 t 5 others
ORIGIN

Query Match 12.5%; Score 254; DB 6; Length 266;
Best Local Similarity 98.4%; Pred. No. 0.00e+00;
Matches 250; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 9 AGCACATTAACCATTTCTATAGCATGATTTCTTCAAGTAAAGGCAAGATATAAATYT 68
|||||
QY 1744 AGCACATTAACCATTTCTATAGCATGATTTCTTCAAGTAAAGGCAAGATATAAATTT 1803
|||||
Db 69 TATAATGACTTGAGTACTTTAAGCCTTGTTTAAACATTTCTTACTTAACCTTTTGCAA 128
|||||
QY 1804 TATAATGACTTGAGTACTTTAAGCCTTGTTTAAACATTTCTTACTTAACCTTTTGCAA 1863
|||||
Db 129 TTAACCCATTGTAGCTTACCTGTAATATACATAGTAGTTTACCTTTAAAGTTGTAAA 188
|||||
QY 1864 TTAACCCATTGTAGCTTACCTGTAATATACATAGTAGTTTACCTTTAAAGTTGTAAA 1923
|||||
Db 189 ATATTGCTTTAACCACACTGTAAATATTTTCAGATAAACATTATATCTTGTATATAAC 248
|||||
QY 1924 ATATTGCTTTAACCACACTGTAAATATTTTCAGATAAACATTATATCTTGTATATAAC 1983
|||||
Db 249 TTTMCACWCYGT 262
|||||
QY 1984 TTTACATCTGTTT 1997
|||||

Search completed: Thu Nov 5 12:14:26 1998
Job time : 4598 secs.

Dec 9 06:58

translate.seq

1

```
; Sequence 2, Application US/08848439
; GENERAL INFORMATION:
; APPLICANT: LaVALLIE, EDWARD
; APPLICANT: RACIE, LISA
; TITLE OF INVENTION: HUMAN SDF-5 PROTEIN AND COMPOSITIONS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CAMBRIDGE PARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,439
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, STEVEN R.
; REGISTRATION NUMBER: 32,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8260
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; Reverse Translation from the peptide US-08-848-439-2.
; Note: the original peptide US-08-848-439-2 contained at least one of
; the residues Arg, Ile, Leu, or Ser. The nucleic acid sequence
; thus contains ambiguous bases which may translate into amino acids
; other than the original amino acids.
848439-2-trans
ATGTCNARGCNCNGGNWNYTNYTNTTYTNGCNWNSCAYTGTYGTYTNGGNWNSGNCNMNG
GNYTNTTYTNGCNARCNGAYTYTWSNTAYARMGNWSNAAYYTYAARCCNATCCNGCNAAYYT
NCARYTNTGYCAGGNATGARTAYCARAAATGMYTNCNAAYYTNTYNGCNAYGARACNATGAAR
GARTNYTNGARGCNGGNCNTGCATCCNTGATCCNTGATCCNTGATCCNTGATCCNTGATCCNT
TYTNTGYSNYTNTTYGNCNGTNTGYTNGAYGAYTNGAYGARACNATNCARCCNTGYCAYWSNYT
NTGYTNCARGTNAARGAYMNTGYGNCNGTNTGATCCNTGATCCNTGATCCNTGATCCNTGATCCNT
CARTGYGAYMNTTYCCNARGCAYAYGAYTNTGYATNCNTYNGCNWNSWNGAYCAYTNTYTCNG
CNACNGARGCNCNAAARTNTGYCARGCNTGYAARAAYARAAYGAYGAYGAYATNATGGA
RACNYTNTGYAARAAYGAYTYTGYCNATNAARATNAARATNAACNTAYATNAAYMGNAYACN
AARATNATNTNGARACNAAARWSNARACNATNTAYAARYTNAAYGGNGTWSNGARMGNYTNAARA
ARWSNGTNTNTGGYTNAARAYWSNYTNCARTGYACNTGYCARGARATCAAYGAYATNAAYGCCNTA
YYTNGTNAATGGNCARACARGGNGGNGARYTNGTNAACNWSNGTNAARMGNTGGCARAARGNCAR
MNGARTTYAARMGNATWSNMGNSNATNMGNAARYTNCARTGY1
; Sequence 3, Application US/08848439
; GENERAL INFORMATION:
; APPLICANT: LaVALLIE, EDWARD
; APPLICANT: RACIE, LISA
; TITLE OF INVENTION: HUMAN SDF-5 PROTEIN AND COMPOSITIONS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CAMBRIDGE PARK DRIVE
; CITY: CAMBRIDGE
```

Dec 9 06:58

translate.seq

2

```
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,439
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, STEVEN R.
; REGISTRATION NUMBER: 32,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8260
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; Reverse Translation from the peptide US-08-848-439-3.
; Note: the original peptide US-08-848-439-3 contained at least one of
; the residues Arg, Ile, Leu, or Ser. The nucleic acid sequence
; thus contains ambiguous bases which may translate into amino acids
; other than the original amino acids.
848439-3-trans
WSNGCNMGNGNYTNTTYTNGCNARCNGAYTYTWSNTAYARMGNWSNAAYYTYAARCCNATNC
CNGCNAAYYTNCARYTNTGYCAGGNATNGARTAYCARAAATGMYTNCNAAYYTNTYNGCNAYGA
RACNATGAARGARTNYTNGARCARGCNGGNCNTGGATNCNTYNGTATGAARCACTGYCAYCCNGAY
ACNARAARTTYTNTGYWSNYTNTTYGNCNGTNTGYTNGAYGAYTNGAYGARACNATNCARCCNT
GYCAYWSNYTNTGYTNCARGTNAARGAYMNTGYGNCNGTNTGATCCNTGATCCNTGATCCNTGATCCNT
NGAYATGYTNGARTGYCAYMNTTYCCNARGAYAYGAYTNTGYATNCNTYNGCNWNSWNGAYCAY
YTNTNCCNACNARGARGCNCNAAARTNTGYCARGCNTGYAARAAYARAAYGAYGAYGAYAAAYG
AYATNATGGARACNTNTGYAARAAYGAYTYTGYCNATNAARATNAARATNAACNTAYATNA
YMGNGAYACNAAARATNATNTNGARACNAAARWSNARACNATNTAYAARYTNAAYGGNGTWSNGARMGN
GAYTYAARAARWSNGTNTYNTGGYTNAARGAYWSNYTNCARTGYACNTGYCARGARATCAAYGAYATNA
AYGNCNTAYTNTGNTATGGNCARAARARGGNGGNGARYTNGTNAACNWSNGTNAARMGNTGGCA
RAARGCNCARMGNCARTTYAARMGNATWSNMGNSNATNMGNAARYTNCARTGY1
```


373 ANKRTCYITNACTYGNACRCANAPNWSRTGRCANGYTGATNGTGYTCRTCNARRTCRT 314

Db 465 yttaythaaymgdytbaengngntay-ytbgmgaayt 499
: : l : l : : : : :
Cp 313 CNARRCANAENGNGCRAANARNWRCANARRAAYT 278

RESULT 12
ID N50029 standard; DNA; 501 BP.
AC N50029;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 446.

ID N50029 standard; DNA; 501 BP.
AC N50029;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 446.
KW Antiviral; cell growth regulator; immune system regulator;

AC N50029;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 446.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.

DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFN γ 446.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; *ss*.
OS *Homo sapiens*.

DE	Sequence encoding new modified human beta interferon polypeptides
DE	IFNX 446.
KW	Antiviral; cell growth regulator; immune system regulator;
KW	antiproliferative; ss.
OS	Homo sapiens.
FH	Key Location/Qualifiers

DE	IFNX 446.
KW	Antiviral; cell growth regulator; immune system regulator;
KW	antiproliferative; ss.
OS	Homo sapiens.
FH	Key
FT	CDS
	Location/Qualifiers
	1..501

KW	antiproliferative; ss.	
KW	Antiviral; cell growth regulator; immune system regulator;	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	1..501
FT	/*tag= a	

KW	antiproliferative; ss.
OS	Homo sapiens.
FH	Key
FT	CDS
FT	Location/Qualifiers
PN	1..501
FT	/*tag= a
PN	EP-163993-A.
FT	13 579 1995

OS	Homo sapiens.
US	Key
FH	Location/Qualifiers
FT	CDS
FT	/*tag= a EP-163993-A. 11-DEC-1985. 17 MAY 1985. 105750
PN	
PD	
PE	

FT	FT	/*tag= a	
PN	EP-16393-A.		
PD	11-DEC-1985.		
PF	17-MAY-1985.	105750.	
PP	17-MAY-1984.	CB-012564	

PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PT Bell LD, Boselev PG, Porter AG;

PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI; 85-311944/50.
DR P-PSDB; P50028.

PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI; 85-311944/50.
DR P-PSDB; P50028.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,

DR GPI; 85-311944/50.
DR P-PSDB; P50028.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2g, page 38; 71pp; English.

PI New modified human beta interferon polypeptide(s) - prepa. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2g, page 38; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinity
for cell surface receptors allowing selective + specific, thou

PS Claim 28; Chart 2g, page 38; 71pp; English.
CC Compared with interferon beta prep'd. by recombinant methods, the
CC INFs of the invention are more active and have different affinity
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against micro-
CC biocides during synthesis and better *in vitro* solubility and

CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts
S0 Sequence 501 BP: 112 A: 31 C: 69 G: 79T:

CC have higher therapeutic index; improved stability against microbial breakdown during synthesis; and better in vivo solubility and stability. They are also easier to recover from incubation mixts

SQ Sequence 501 BP; 112 A; 31 C; 69 G; 79 T;

Query Match 6.8%; Score 50; DB 3; Length 501;

CC stability. They are also easier to recover from incubation mixts

SQ Sequence 501 BP; 112 A; 31 C; 69 G; 79 T;

Query Match 6.8%; Score 50; DB 3; Length 501;

Best Local Similarity 16.0%; Pred. No. 1.54e-04;

Matches 72; Conservative 133; Mismatches 242; Indels 4; Ga

Query Match 6.8%; Score 50; DB 3; Length 501;
Best Local Similarity 16.0%; Pred.No. 1.54e-04;
Matches 72; Conservative 133; Mismatches 242; Indels 4; Gaps 0

Db 53 araarybybtbggcarytbaayvggnmgdytbgartaytgyytbaargaymgdatgaayt
|| : ||| : | : : | : : ||| : || : ||| : | : : | : | :
Cp 728 ARNWSRTCYTTNARCCARNACNWSYTTYTTTNAARTCNCKYTCNWSNACNCCTTNARY

[illegible]

Cp 728 ARNWSRSGYTTNARCCANARNACNWSYTTYTTNARTCNCKYTCNWSNACNCCRTTNARY
 Db 113 tyggcnathccngargarathaarcaargcngncarcattycaraargaygcngcnytbab
 Cp 653 TTRATANATGTYTNNWSYTTNGTYTCNARNATNATYTTNGTRTCNCKRTTNATRTANGTN

Db 113 tyggnathccngargarathaarcargncarcarttycaraargaycngcnybtba
l:: : l: : : | l: | : : :
Cp 663 tttgtatnctytnwsttngtytcnarnaTnatttngctrctcnckrtTTNATRTANGTN
Dd 173 cnathcaygaratgytbcaarayathhttygcnahtltymgdcardgaywnnnwnnwnaeng
|::: | | :: | : |: |: | | :: :

[illegible]

173 cnathaygaratgytbcaraayathhttygcnahthttymgdcargaywnnnwnnnwnnacn
608 ATYCYTNACYTNAITYTNARNGCRAARTCRITYTTRCANARNCTYTCCATNA TRTCR
233 gntggaaygaracnathgtngaraaaytbytbgcnaaygntntaycaycarathaaycayy

Cp 608 ATTCCTTACATTTATTTTAAAGGCGGAAATCTTTTTTCAGAAAGGTTTTCCAGATATATCC
 Db 233 gntggaaygaracnathgtngaraaaytbytbgnaaaygntntaycaycarathaaycay
 : :: : |::| : :||:| |::| : :| : :
 Cp 548 TTATCTGCTCTTCTTTTATTTTCANGCVTCRCANACYTTNGGNGCYTCYTCNGTNGCN
 Db 293 tbaaracnctnbybqargara--arytbqaraa--rgargayttyacnmgdgnaaryvtba

548 TTRTCRTRCTRTYTTTTRTYTTRCANGCYTCRCANACYTTNGGNCYTCYTCNGTNGCN

Cp

Db

293 tbaaracngtnytcbgargara--arytbqaraa-rgargayttyacnmgddgnaarytba

Cn

488 GGNARNABRTCTRCNTGNSWNGCNA BNGCNATBCANABRTCTRTCTTCYTTGNGGRRANCKR

[illegible]

Cp 488 GGNARNARRTGRTCNWSNWSNGCVARNGGNATRCANARRCTTRTCTYTGNGGGAANKCR

Cp	488	GGVARNARRTGRTCNWSNNSGNCNARNGGNATRCANARRTCRTTRTCYTGNGGGRANCKR	429
Db	350	tgwnnnnybtbcayytbaarmgdtaytayggnmgdathybtbcaytaytbaargcnaarg	409
Cp	428	TCRCAYTCNARCARTTCNGGCCANGGAAACCAACCCRAANGCNWSCATNACNGGNGCRCANCKR	369
Db	410	artaywnncaytgygcntggacnathgtmngdgtngarathybtmgdaaytyttayttaa	469
Cp	368	TCYTTNACYTGNACRCANARNWSRTGRCANGGYTGNAATNGTYTCRTCNARRTCRTCNARR	309
Db	470	thaaymgdytbacnggntay-ytbmgdaayt	499
Cp	308	CANACNGGNGCRAANARNWSRCANARRAAAT	278

RESULT 15
ID N50025 standard; DNA; 501 BP.
AC N50025;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 418.

KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
1		

FT	CDS	1..501
----	-----	--------

FT /*tag= a

PN EP-163993-A.

PD 11-DEC-1985.

PF 17-MAY-1985; 105750.

PR 17-MAY-1984; GB-012564.

PA (SEAR) SEARLE G D & CO.

PI Bell LD, Boseley PG, Porter AG;

DR WPI; 85-311944/50.

DR P-PSDB; P50024.

PT New modified human beta interferon polypeptide(s) - prepd. by

PT plasmid transformed bacteria, with improved antiviral,

PT anti-proliferative and immune regulating actions

PS Claim 28; Chart 2c, page 34; 71pp; English.

CC Compared with interferon beta prepd. by recombinant methods, the

CC INFs of the invention are more active and have different affinities

CC for cell surface receptors (allowing selective targeting); they

CC have higher therapeutic index; improved stability against microbial

CC breakdown during synthesis; and better in vivo solubility and

CC stability. They are also easier to recover from incubation mixts.

SQ	Sequence	501 BP;	112 A;	30 C;	69 G;	85 T;

Query Match

Best Local Similarity 15.7%; Pred. No. 2.99e-04;

Matches 71; Conservative 134; Mismatches 242; Indels 4; Gaps 3;

Db 53 araaarybtbtggcarytbaayggnmgdytbgartaytggytbaatgaymgdatgaayt 112

Db 113 tygayathcncatgargarttygayggnaycartyccaraargaygcngcnytba 172

Db 173 cnathtaygaratgytbcataayathtttygcnathttymgdcatgywnnwnnnaeng 232

Db 233 qntqqaayqaracnathqtngaraayvtbvtbqcnaayqntaycaycarathaaycayv 292

[illegible]

Search completed: Tue Dec 9 09:10:39 1997

Job time : 282 secs.